Working Paper Series

Reassessing the Gender Wage Gap: Does Labour Force Attachment Really Matter? Evidence from Matched Labour Force and Biographical Surveys in Madagascar

Christophe Nordman
François Roubaud

ECINEQ WP 2005 – 16

ECINEQ 2005-16
December 2005

www.ecineq.org

Reassessing the Gender Wage Gap: Does Labour Force Attachment Really Matter?

Evidence from Matched Labour Force and Biographical Surveys in Madagascar*

Christophe Nordman and François Roubaud†
DIAL, IRD-Paris

December 2005

Abstract

Assessing gender inequalities has become one of the key issues of the new international poverty reduction strategies implemented in most LDCs in the past few years. It has been argued that differences in labour force attachment across gender are important to explain the extent of the gender earnings gap. However, measures of women's professional experience are particularly prone to errors given discontinuity in labour market participation. For instance, the classical Mincerian approach, where potential experience is used as a proxy for actual experience due to lack of appropriate data, has its limits in estimating the true returns to human capital. Such biases in the estimates cannot be ignored since the returns to human capital are used in the standard decomposition techniques to measure the extent of gender-based wage discrimination. By matching two original surveys conducted in Madagascar in 1998 - a labour force survey and a biographical survey - we built a unique dataset that enabled us to combine the original information gathered from each of them, particularly the earnings from current employment and the entire professional trajectories. Our results lead to an upward reappraisal of returns to experience, as potential experience always exceeds actual experience, for both males and females. In addition, controlling for further qualitative aspects of labour force attachment, we obtain a significant increase in the portion of the gender gap explained by observable characteristics, while the differences in average actual experience across sexes lead to markedly different estimates of the fraction of the gender earnings gap explained by experience.

Keywords: gender wage gap, returns to human capital, labour force participation, biographical data, Madagascar
JEL Classification: J24, J31, O12

---

* We gratefully acknowledge valuable comments from participants in a seminar at DIAL in Paris, in the Minnesota International Economic Development Conference in Minneapolis, in the Journées de l'AFSE in Clermont-Ferrand, in the Annual Conference of the European Society for Population Economics (ESPE) in Paris and in the Meeting of the Society for the Study of Economic Inequality (ECINEQ) in Palma de Mallorca. Special thanks go to Ragui Assaad, Denis Cogneau, Philippe de Vreyer and François-Charles Wolff for helpful remarks.Usual disclaimers apply.
† **Contact details**: DIAL, IRD-Paris. 4 rue d'Enghien, 75010 Paris, France. Emails: nordman@dial.prd.fr ; roubaud@dial.prd.fr

1. Introduction

Returns to human capital have always been considered dominant explanations for labour compensation. In individual wage equations, researchers typically include human capital and skills through regressors describing the worker's schooling and labour market experience (Mincer, 1993; Card, 1999). This is particularly important for developing countries where the returns to education are expected to be higher (Sahn and Alderman, 1988; Hoddinott, 1996; Behrman, 1999; Schultz, 2004). However, before the 1980s, human capital accumulated on-the-job was hardly properly measured.[1] The recommended estimate consisted in using the time spent in certain circumstances, i.e. in the firm or the workplace. Since measures of actual experience were not available when the major empirical developments of the original theory emerged, estimates were established using potential experience, calculated as age minus years of schooling minus age on entering school. Refinements were proposed later, as new surveys became available providing more detailed information about the time that individuals had actually devoted to their principal employment. Hence, Mincer and Jovanovic (1981) introduced the workers' tenure in firms to take into account the return to specific training received. The time elapsed in the labour market is assumed to reflect the accumulation of general human capital. The remuneration of experience and tenure therefore represents the return to human capital accumulated on the job. The longitudinal data available today distinguishes more accurately between these two measures and enriches the information used in empirical studies. It is therefore not only possible to calculate more or less exactly the time that an individual has dedicated to work, but also to isolate the experience acquired in various industries and/or jobs. Nowadays, studies using this type of measure are frequent in developed countries, too frequent indeed to be detailed here (see for example Kim and Polachek, 1994; Light and Ureta, 1995 or Myck and Paull, 2004).

These issues are of a great importance in assessing the extent of gender inequalities in labour markets. In industrialised countries, many attempts have been made to estimate the extent to which the average gender wage gap is due to differences in human capital attributes such as schooling and work experience, versus differences between genders in wages paid for given attributes (Blau and Kahn, 2000). From the literature on this issue, less than half of the gap can be explained by factors such as differences in years of schooling and experience and

[1] Mincer (1974) had already admitted that the representation of post-school investments was the weak point of the theoretical architecture of his model: "[…] *the most important and urgent task is to refine the specification of the post-school investment category […] to include details (variables) on a number of forms of investment in human capital*" (Mincer, 1974, p. 143).

tenure. However, it has been shown that missing or imprecise data on these human capital factors can result in serious biases in the calculation of the discrimination component resulting from Mincerian wage equations (Stanley and Jarrell, 1998; Weichselbaumer and Winter-Ebmer, 2005).

In fact, measures of women's professional experience are particularly prone to errors given discontinuity in labour market participation. Often age or the Mincer measure of potential experience are still used as a proxy for the acquisition of general human capital or for work experience. Potential experience may be a good approximation of true experience for men with high labour force attachment, but is a poor proxy for less attached individuals, especially for women or minority groups as they have a greater likelihood of interrupting their professional activities (Altonji and Blank, 1999; Antecol and Bedard, 2004). Proxy measures tend to overstate women's actual work experience by not accounting for interruptions related to parenting (that is, complete withdrawals from the work force) or, for instance, for any restrictions on the number of hours worked per week. Furthermore, empirical studies have revealed that experience before an interruption has a lower return than experience after an interruption and that women who interrupt their careers generally receive less wage growth prior to the interruption (see for instance Mincer and Polachek, 1974; Sandell and Shapiro, 1980; Mincer and Ofek, 1982; Adair *et al*., 2002). Hence, the coefficient of experience in the wage equation, but also the coefficient of education, may be systematically biased, notably for women.[2] Such biases in the estimates cannot be ignored since the returns to human capital are used in the standard decomposition techniques for gender wage gaps, and therefore to measure the extent of gender-based wage discrimination (Blinder, 1973; Oaxaca, 1973). Authors have in fact argued that these measurement errors can amplify the impact attributed to pure discrimination (the unexplained part of the wage decomposition), to the detriment of the component relating to observed differences in individual characteristics between men and women (Stanley and Jarell, 1998; Weichselbaumer and Winter-Ebmer, 2005).

In the case of Africa, there is very little known about the gender wage gap (Appleton, Hoddinott and Krishnan, 1999). From Weichselbaumer and Winter-Ebmer (2005)'s recent meta-analysis on this issue, we can evaluate that only 3% of the studies on gender wage gap stem from African data out of all the empirical literature since the 1960s. From the existing

---

[2] Indeed, it can be shown that underestimating the return to experience can lead in turn to underestimating the return to education if experience and education are substitutes (Dougherty, 2003).

literature,[3] there is however a wide consensus on the presence of important inequalities between men and women, both for salaried and self-employed workers. For instance, in Guinea, Glick and Sahn (1997) find that differences in characteristics account for 45% of the male-female gap in earnings from self-employment and 25% of the differences in earnings from public-sector employment while, in the private sector, women actually earn more than men.

Armitage and Sabot (1991) also found that such gender inequality exists in the public sector of Tanzania but observed no gender discrimination in Kenya's labour market. The latter result is true both for the public and private sectors of the Kenyan economy. Similarly, Glewwe (1990) found no wage discrimination against women in Ghana. On the contrary, females seem better off than males in the public sector. More recently, Siphambe and Thokweng-Bakwena (2001), using data from the 1995-1996 Labour Force Survey in Botswana, show that in the public sector most of the wage gap is due to differences in characteristics between men and women and not to discrimination on the basis of rewards to those characteristics. On the other hand, in the private sector, most of the wage gap is due to discrimination. Likewise, in Uganda and Côte d'Ivoire, Appleton *et al.* (1999) find evidence that the public sector practises less wage discrimination than the private sector. However, from their study on Côte d'Ivoire, Ethiopia and Uganda, they finally conclude that there is no common cross-country pattern in the relative magnitudes of the gender wage gaps in the public and private sectors.[4]

Generally speaking, results from these case studies in Africa suggest the importance of sectoral choices, but also of workers' job status, for analysing differences of wage determination between the sexes. Moreover, in many African countries, the predominance of the informal activity as well as the decreasing role of the public sector in providing stable jobs for qualified workers may give rise to significant selection effects at labour market entry. Hence, labour market segmentation and professional segregation are probably the rule rather than the exception (Cogneau, 1999). This professional segregation may reflect discriminatory practices (sexist recruitment methods, stereotypes and prejudice against women, etc.) based on what Bourdieu (1998) calls “male domination”, which prevents women from having access to certain well-paid segments or professions.

[3] See, notably, Glewwe (1990) for Ghana; Cohen and House (1993) for Sudan; Milne and Neitzert (1994) and Agesa (1999) for Kenya; Glick and Sahn (1997) for Guinea; Armitage and Sabot (1991) for Kenya and Tanzania; Appleton, Hoddinott and Krishnan (1999) for Uganda, Côte d'Ivoire and Ethiopia; Isemonger and Roberts (1999) for South Africa; Siphambe and Thokweng-Bakwena (2001) for Botswana and Nordman (2004) for Morocco and Tunisia.

[4] In Uganda, the authors find that the wage gaps in the public and private sector are comparable. In Ethiopia, there is a much wider gap in the private sector than in the public sector. In Côte d'Ivoire, the reverse is true.

In Madagascar, the only study we are aware of is that of Nicita and Razzaz (2003). The authors investigate the gender wage gap in relation to an analysis of the growing potential of a particular economic sector, the textile industry. From their earnings differential decomposition (Oaxaca, 1973), they show that both the endowments and the unexplained part of the wage difference favour male workers, although the latter dominates the former.[5] Second, education and potential experience (measured by age) are similarly important in determining the wage differential. Third, level of education and being resident in urban Antananarivo slightly reduce the unexplained part of the wage differential. However, no general conclusion on Madagascar can be drawn from their analysis as it only concerns one particular manufacturing sector. Another limitation of their study is that, as a result of lack of information, they proxy total experience by age and include very few and imprecise regressors in their wage equations by sexes. As pointed out earlier, this may have serious consequences on the extent to which gender wage discrimination is appreciated (upward biased) because the unexplained gender wage gap can be attributed in this case to the specification error in the original wage regression (i.e. unaccounted characteristics remain correlated with the unexplained component of the gender wage differential).

Enhancing the gender gap literature on developing countries, especially on the poorest ones, is crucial for several reasons. First, as mentioned earlier, there are manifest shortcomings of studies on African countries, particularly due to the shortage of available information (Bennell, 1996). Second, gender inequality is likely to be greater while markets do not function efficiently and the States lack the resources for introducing corrective policies. Third, understanding the roots of inequalities between the sexes and reducing gender gaps have a central place in term of policies in these countries. For example, under the Poverty Reduction Strategy Paper (PRSP) initiative that concerns over sixty of the world's poorest countries, policies designed to counter gender discrimination are among the most often recommended solutions to combat poverty (Cling, Razafindrakoto and Roubaud, 2003): Goal 3 of the Millennium Development Goals (MDG) is aimed at reducing gender inequalities. Fourth, the above-mentioned problems of labour market attachment for females are even greater than in developed countries. For instance, in the Madagascan case that interests us here, the continual deterioration of the labour market as well as the partial freeze on public sector recruitment from the mid-1980s may have accentuated the circumstances (i.e. labour market entry and

[5] In 1999, the gross unadjusted wage differential is about 51% in favour of males. The results of the decomposition attribute about 14% to differences in endowments. The unexplained part accounts for about 59% of the wage differential, while the remaining 27% is due to selectivity.

exit) that would give rise to measurement errors in labour force attachment, especially for women, for whom the decrease in jobs in the public sector was particularly significant.

In this article, we will cast new light on these issues by using a series of first hand surveys carried out in 1998 in the capital of Madagascar, Antananarivo, under the supervision of one of the authors. The approach consists in matching a labour force survey and a biographical survey in a view to obtaining detailed information on complete professional and family trajectories for a representative sample of the urban population. In particular, we are able to combine the income from current employment, taken from the first survey, with the individuals' actual experience (length and type of jobs occupied, periods of inactivity, unemployment, work interruptions, etc.) over their entire life span, taken from the second. As far as we know, this is the first such attempt at a detailed study of this sort in Africa, which was inaccessible until now due to the lack of appropriate data. The uniqueness of our data enables us to control the effects of selection on labour market entries, and to differentiate between men and women in this respect. We thus propose different decompositions of the gender earnings gap that take into account: (1) the effects of selection relating to labour force and sectoral choices (public, formal private and informal/self-employed sectors) and (2) alternatives to the standard methods for measuring workers' labour force attachment.

Our results lead to an upward reappraisal of returns to experience, as potential experience always exceeds actual experience, for both males and females. In addition, controlling for further qualitative aspects of labour force attachment, we obtain a significant increase in the portion of the gender gap explained by observable characteristics, while the differences in average actual experience across sexes lead to markedly different estimates of the fraction of the gender earnings gap explained by experience.

The rest of the paper is divided as follows. Section 2 briefly presents the two datasets, the background of the Madagascan labour market and some descriptive statistics. Section 3 discusses the main econometric methods for assessing the gender gap: earning functions and gender wage decompositions. In section 4 we comment on the econometric results. Finally, in section 5, we draw together the main findings and conclude.

2. Data, Madagascan context and descriptive statistics

2.1 The data: matching labour force and biographical surveys

The data used here has been obtained by matching two original surveys conducted in Madagascar in 1998 by the National Institute of Statistics (INSTAT) as part of the MADIO project (MADIO, 1998, 1999; Roubaud, 2000):

- the first, a labour force survey, was designed to collect detailed information on employment, unemployment, income and working conditions in the Madagascan capital, Antananarivo;
- the second, a biographical survey, followed the trajectories of a representative sample of Tananarivians in three different fields: migratory and residential trajectory, family and marital trajectory and schooling and professional trajectory.

The joint use of these two surveys offers three key advantages for our study. First, this type of survey, whether on labour force or on individual trajectories, is extremely rare in the African context. Second, the data is of a far higher standard than that usually collected in household surveys in Africa (Rakotomanana *et al.*, 2003). Finally, the fact that the sample used in the biographical survey was a sub-sample of the labour force survey means that the two surveys can be matched on an individual level, thereby enabling us to combine the original information gathered for each of them, particularly the earnings from current employment in the labour force survey and the entire social and professional trajectories in the biographical survey (individual's household characteristics, employment, unemployment, inactivity spells).

The labour force survey corresponds to the first phase of the *1-2-3 Survey*, on employment, the informal sector and consumption, carried out in Madagascar by the National Institute of Statistics (INSTAT) since 1995, as in a number of others developing countries, in Africa and in Latin America (Razafindrakoto and Roubaud, 2003; Cling *et al.*, 2005). The sample, drawn from a stratified two-stage area-based survey plan, is representative of all ordinary households in Antananarivo. In 1998, of the 3,002 households questioned, we counted 10,081 people of working age, of whom 5,822 individuals were active wage earners, 361 unemployed and 3,998 inactive. The definitions used (activity, unemployment, etc.) follow the international standards recommended by the ILO in this respect (Hussmann, Mehran and Verma, 1990). For all those in work, we have a comprehensive set of data on the job characteristics. Special attention is given to capturing income derived from work. In the 1998 survey, out of a total of 5,298 active wage-earners, 3,445 declared their actual income and 1,853 declared their income bracket (expressed in multiples of the current minimum wage); only 13 individuals refused to provide information on their income, which is in itself an indicator of the quality of the survey. The survey also provides an estimate of the total benefits relating to the job (sundry bonuses, paid holidays, housing, benefits in kind, etc.), whether monetary or non-monetary, which are added to the direct income.[6] We should also point out that, since all the members of the household are interviewed for the survey, we measure the total household income and can also identify each individual's contribution. This variable is particularly interesting when it comes to estimating the individual labour supply, notably depending on the income of other members of the household.

The biographical survey (*Biomad98*) follows on from the one carried out in France in 1981 by the French National Institute of Demographic Studies (Courgeau and Lelièvre, 1992) and in a certain number of African capitals from the end of the 1980s (Dakar, Bamako, Yaoundé, Lomé, Nairobi; see GRAB, 1999; Antoine *et al.*, 2004). These surveys are retrospective, and are aimed at describing different aspects of urban integration processes: access to employment, access to housing, family formation and demographic dynamics. This type of approach helps analyse interactions between family situations and residential and professional trajectories. By introducing a time factor, the biographical surveys can be used as a complement to setting up panel data. Although the retrospective nature of the surveys can impair the quality of the information collected due to memory problems on the part of the respondents, they do have two key advantages: they are not subject to the problem of attrition, which is especially difficult to manage with panels, and they piece together the respondents' entire trajectories.

The *Biomad98* survey addressed three generations of individuals: those born between 1943 and 1952 (aged 45-54 at the time of the survey), between 1953 and 1962 (35-44) and those born between 1963 and 1972 (25-34). 2,403 biographical questionnaires were collected among the individuals identified in the labour force survey, using a “grafting” technique to combine the surveys. In order to obtain a representative sample of the three generations in question, and to enable separate analysis for men and women, the main object of the study, we decided to survey around 400 people for each of the six cohorts concerned. We therefore used a survey plan stratified by generation and by gender.

[6] As is the case in all surveys of this kind, measurement errors are greater for non-salaried workers, particularly in the informal sector. However, phase 2 of the *1-2-3 Survey* (not used in this article), which pieces together all the production accounts and income accounts for informal production units, helped confirm that the income declared in the employment survey was in fact coherent.

The matched data allows us to construct several measures of actual (rather than potential) work experience: experience off the incumbent firm or main employment, years of tenure with the current employment, in the main occupation and in the main profession.[7] Potential experience is simply age minus years of education minus six. Actual experience is measured as months worked at the time of the *Biomad98* interview and is converted into years of experience. Other labour force attachment measures include: the time spent out of the labour force (inactivity), unemployment periods, as well as the number of work interruptions over individuals' lives, from the end of school until the date of the interview (or from the age of six if they have zero years of school attendance). This last variable is incremented by one from zero each time a spell of declared work has been interrupted by either a period of education, inactivity or unemployment. Non-working time (unemployment plus spells out of the labour force) is similarly accumulated from the age of six onwards and is calculated in years. In the rest of the paper, all these measures shall be referred to as 'labour force attachment variables' (LFAVs).

In the data, the labour supply or paid work participation has been defined as individuals having worked at least one hour during the reference week and reporting positive earnings at the time of the interview. For those individuals who have declared positive earnings (1,928 out of 2,403 individuals), we have identified three institutional sectors of paid work participation: public wage employment, formal private wage employment and self-employment or informal sector, defined as those working in production units that are not registered or do not publish accounts.

Finally, matching these two sources of information allowed us to build a unique dataset containing biographical-type information on the individuals' socio-economic characteristics together with a series of variables on their activity, labour incomes and job characteristics. The biographical data, spanning individuals' entire professional careers, provides relevant information that can be used to improve standard measures of human capital.

[7] For instance, we can distinguish the length of different types of work experience: the time elapsed with the same employer (tenure strictly speaking), the time spent in the same occupation (taking into account the fact that workers may have two different and successive occupations with the same employer), as well as the years of experience in what individuals consider as their main "profession" (e.g. a carpenter who has practised his or her main duties in different workshops or firms).

2.2 Madagascan background and descriptive analysis

Over the past fifteen years or so, Madagascar, one of the poorest countries in the world, has embarked on a process of economic liberalisation, similarly to many African countries undergoing structural adjustment. Over the long term, Madagascar is distinguished by a constant decline in household living standards, which in 1996 reached the lowest point since independence. From the mid-1990s, the reform process began to bear fruit. In 1997, growth in GDP per capita was slightly positive (1%), for the first time in many years. This historic shift then accelerated, with growth reaching 4% in 2001. The contested Presidential elections in December 2001, followed by the open political crisis that continued throughout the first six months of 2002, jeopardized economic improvements, and living standards once again fell sharply (Roubaud, 2002). Since then, the country has been trying as best it can to recover.

In 1998, the period referred to in this article, there had already been a very significant recovery in urban areas, especially in the capital. In three years, from 1995 to 1998, the average real labour income grew by 35% and the median income by 51%. The side-effects of growth had a very positive impact on the labour market: increase in schooling, decrease in child labour, slight decrease in unemployment, which is structurally low, but above all an end to the informal sector's domination of the labour market and a massive drop in underemployment and poverty. The incidence of extreme poverty (with the poverty line at US$ 1 in PPP) fell from 39% to 28%. In terms of gender, women's activity rate fell, corresponding to the withdrawal from the labour market of large numbers of women who had been forced to work to provide additional income for their households during a severe crisis. At the same time, the income differential between men and women was reduced (Razafindrakoto and Roubaud, 1999).

Despite improvements in the situation, the three years of recovery were not enough to erase several decades of continual deterioration in the labour market. In the long-term perspective that interests us here, the main characteristic of labour market evolution was the partial freeze on public sector recruitment from the mid-1980s, which went hand in hand with a fall in the numbers of wage-earners and an underlying rise in job precariousness. The decrease in jobs in the public sector was particularly significant for women (Antoine *et al.*, 2000). In our data, the patterns of participation and sectoral distributions differ sharply across gender. The participation rate is much lower for women (95% against 78%), while unemployment is not significantly different by sex (3%). Among occupied workers (92% and 75% of males and females respectively), women are concentrated in low quality jobs in the informal sector, which represents 44% and 55% of occupied males and females respectively. Consequently, their presence in the public sector is 8 points lower than for men (25% against 17%). Men and women bring also different work experience to the labour market.[8] The Mincer proxy for potential work experience shows little difference in the work experience of men and women (22.6 and 24.0 years), as the average age is the same, while the average years of education (successfully completed or not) are about one year lower for women. A different story emerges when actual experience is applied. The average actual work experience is 20.5 years for men compared with 17.1 years for women.

Differences in earnings and human capital across gender and generation

Variables	Males	Females	Difference
	Mean	Mean	
Hourly earnings*	2.20	1.52	0.68
Generation 1963-1972			
Hourly earnings	1.42	0.94	0.47
Years of schooling successfully completed	9.60	9.04	0.57
Years of potential experience	10.75	11.71	-0.96
Years of actual experience	9.22	7.53	1.69
Generation 1953-1962			
Hourly earnings	2.32	1.55	0.77
Years of schooling successfully completed	9.28	7.75	1.53
Years of potential experience	20.98	23.30	-2.32
Years of actual experience	18.87	14.94	3.93
Generation 1943-1952			
Hourly earnings	2.59	1.63	0.95
Years of schooling successfully completed	8.38	7.12	1.26
Years of potential experience	31.76	33.39	-1.63
Years of actual experience	29.13	21.64	7.49

Sources: *Enquête 1-2-3*, *Phase 1*, *1998*, *Biomad98*, MADIO; authors' calculations.
* : in Madagascan Francs (Fmg).

Disaggregating by cohort gives a more precise view of the biases caused by only taking into account potential experience (Table abeove). The bias is highest for women in the eldest generation. While the difference between potential and actual experience is 4.2 years for the youngest generation of women, it increases to 11.8 for the eldest. For men, the gap is more or less constant across the cohorts (around 2 years). This result is explained by the accelerated demographic transition process in the Madagascan capital. The number of descendants has fallen significantly in the past three decades. For example, at the age of 30, women belonging to the 1943-1952 generation had 3.4 children; at the same age, the intermediate generation only had 2.7, whereas the youngest generation has 1.8. This fall in fertility rates comes from

[8] Descriptive statistics for paid-work participants are shown in Appendix, Table 1.

later first births (at 25, three-quarters of women in the eldest generation had had at least one child, against barely half in the youngest generation), and also from higher intergenesic intervals, for which the median period increased from 37 months to 67 months from the eldest to the youngest generation (Antoine *et al.*, 2000).

3. Econometric methods

3.1 Earnings determination

Earnings functions and correction for selectivity

Traditional gender wage decompositions rely on estimations of Mincer-type earnings functions for men and women. Let the earning function take the form:

$$\ln w_i = \beta x_i + \varepsilon_i \qquad (1)$$

where $\ln w_i$ is the natural logarithm of the observed hourly earnings for individual i, x_i is a vector of observed characteristics, β is a vector of coefficients and ε_i is a disturbance term with an expected value of zero.

We estimate the log earning functions for the pooled sample and, then, separately for males and females and for the different sectors. There is no universally accepted set of conditioning variables that should be included for describing the causes of gender labour market differentials. However, the consensus is that controls for productivity-related factors such as education, experience and tenure, marital status, presence of children, and location of residence[9] should be included. However, it is debatable whether job characteristics, occupation and industry should be taken into account: if employers differentiate between men and women through their tendency to hire into certain occupations, then occupational assignment is an outcome of employer practices rather than an outcome of individual choice or productivity differences.[10] We also incorporate in the earnings functions a dummy for formal training received during the current employment and paid by the employer. More educated workers generally receive more job training (Barron, Berger and Black, 1999), which is indeed the case in our sample. Besides, introducing this variable may help to

[9] In our data, this information is not relevant as all individuals live in Antananarivo and its close outskirts.
[10] Conversely, one can argue that analyses that omit occupation and industry may underestimate the importance of background and choice-based characteristics on labour market outcomes (Altonji and Blank, 1999).

attenuate the effects of unobserved skills of the workers, since more able employees may receive more on-the-job training.

Thanks to the longitudinal information available in the biographical survey, years of labour market experience, that are commonly proxied by potential measures, are refined by using actual measures of experience as well as other labour force attachment variables (LFAVs, see section 2.1) to take into account possible differentiated human capital depreciation (or appreciation) effects (Mincer and Ofek, 1982). Other regressors include dummies on marital, religious and ethnic status, a dummy for the presence of a union in the current job, two dummies for the type of work contract (the reference is no contract), and the number of hours worked per week. Following standard procedure (see Altonji and Blank, 1999), sectoral and occupational dummies are included as independent variables but separately in the earnings decompositions, so as to propose alternative measures of the gender earnings gaps.

Concerns arise over possible sample selection biases in the estimations. Strictly speaking, there are two sources of selectivity bias involved. One arises from the fact that wage-earners are only observed when they work, and not everyone is working. The second comes from the selective decision to engage in public wage employment rather than private wage employment or the informal sector. We use Heckman's two-step procedure to address the first issue. In the first stage, probit estimates of the probability of participation are separately performed for males and females. We then include the appropriate estimated correction term (Inverse Mill's Ratios, IMR) into the second-stage earnings equations, for males and females respectively. The inclusion of the correction term ensures that the OLS gives consistent estimates of the augmented earnings functions.

One way to account for the second issue is to determine whether the returns to characteristics of a wage-earner differ from one institutional sector to another. However, given the over-representation of men in the state sector, the decision to work in a particular sector may not be determined exogenously. Apart from the observed characteristics of women discussed earlier (such as education and experience), it may correlate with unobserved characteristics. We use Lee's two-stage approach to take into account the possible effect of endogenous selection in different sectors on earnings (Lee, 1983). In the first stage, multinomial logit models of individual i's participation in sector j are used to compute the correction terms λ_{ij} from the predicted probabilities P_{ij}. The appropriate correction term is then included in the respective earnings equation as an additional regressor in the second stage.[11] Lee's method has been recently criticised in the literature because it relies upon a strong assumption regarding the joint distribution of the error terms of the equations of interest (see Bourguignon *et al*., 2004). However, the existing alternative methods we tried, such as Dubin and McFadden's, did not appear more efficient given the limited size of our sectoral sub-samples. As showed Bourguignon *et al*. (2004), Lee's method is indeed adapted to limited sample sizes.

A multinomial logit model with four categories is specified. It includes non-participation in paid employment (as the base category), public wage employment, formal private wage employment and self-employment or informal sector. In both Heckman's and Lee's procedures, identification is achieved by including various household variables (mainly drawn from the biographical survey), such as dummies for the status of the individual in the household (household head, head's spouse, head's children, head's parent), the number of children by age categories (aged 0-4, 5-9 and 10-14), the household's income per capita (without the individual's contribution), the inverse of the dependency ratio (number of working individuals divided by the total number of individuals in the household), a material wealth proxy[12], father and spouse's education, spouse's religion and ethnicity, dummies for the status of the individual vis-à-vis his/her housing and whether housing receives electricity. From the biographical data, it is also possible to test whether past events, and particularly their order of occurrence, can influence individuals' situations with respect to the labour market at the time of the interview. For instance, we identified whether individuals were already married before getting their first job and added a dummy in the participation equations. This variable has arguably no theoretical reason to influence earnings determination but may influence employment participation, especially that of women who must balance domestic responsibilities with the need to augment family income.[13] Therefore, it appears as a very good identifying variable for the selection equation since it is uncorrelated to the error term of the earnings equation. This is one way of overcoming the limitation of Heckman's two-step procedure, i.e. to find additional variables that arguably do affect work force participation in the first step but have no direct impact on earnings in the second.[14]

[11] The presence of the additional constructed selectivity correction terms renders the standard errors incorrect. White's standard errors are used to provide asymptotically consistent values.

[12] The sum of the number of house, car, fridge, television, hi-fi, phone, radio and stove.

[13] Theoretically, getting married before having a first job may raise women's opportunity costs to labour market participation and, therefore, their reservation wage. If it is the case, the expected impact of this variable on the probability of being employed at the time of the survey should be negative (with time, their incentives to participate may be less and less important as well as their employability). However, the presence of children soon after a marriage may exert a contradictory effect since children require care and supervision, but they also increase the needs for market goods, so for labour income (see Glick and Sahn, 1997).

[14] The data confirms this assumption.

3.2 Gender wage decomposition techniques

Oaxaca and Neumark's traditional decompositions

The most common approach to identifying sources of gender wage gaps is the Oaxaca-Blinder decomposition (e.g., Oaxaca, 1973; Blinder, 1973). Two separate standard Mincerian log wage equations are estimated for males and females. The Oaxaca decomposition is:

$$\overline{\ln w_m} - \overline{\ln w_f} = \beta_m(\bar{x}_m - \bar{x}_f) + (\beta_m - \beta_f)\bar{x}_f \tag{2}$$

where w_m and w_f are the means of males and females' wages, respectively; x_m and x_f are vectors containing the respective means of the independent variables for males and females; and β_m and β_f are the estimated coefficients. The first term on the right hand side captures the wage differential due to different characteristics of males and females. The second term is the wage gap attributable to different returns to those characteristics or coefficients.

In equation (2), the male wage structure is taken as the non-discriminatory benchmark. It can be argued that, under discrimination, males are paid competitive wages but females are underpaid. If this is the case, the male coefficients should be taken as the non-discriminatory wage structure. Conversely, if employers pay females competitive wages but pay males more (nepotism), then the female coefficients should be used as the non-discriminatory wage structure. Therefore, the issue is how to determine the wage structure β^* that would prevail in the absence of discrimination. This choice poses the well-known index number problem given that we could use either the male or the female wage structure as the non-discriminatory benchmark. While *a priori* there is no preferable alternative, the decomposition can be quite sensitive to the selection made. If we let:

$$\beta^* = \Omega\beta_m + (I - \Omega)\beta_f$$

where Ω is a weighting matrix and *I* is the identity matrix, then any assumption regarding β^* can be seen as an assumption regarding Ω. The literature has proposed different weighting schemes to deal with the underlying index problem: first, Oaxaca (1973) proposes either the current male wage structure as β^*, i.e. Ω=*I*, or the current female wage structure, Ω=0, suggesting that the result would bracket the "true" non-discriminatory wage structure. Reimers (1983) implements a methodology that is equivalent to Ω=0.5 *I*. In other words, identical weights are assigned to both men and women. Cotton (1988) argues that the non-discriminatory structure should approach the structure that holds for the larger group. In the context of sex discrimination, such weighting structure implies an $\Omega = I_m I$ where I_m is the fraction of males in the sample.

Neumark (1988) proposes a general decomposition of the gender wage differential:

$$\overline{\ln w_m} - \overline{\ln w_f} = \beta^*(\bar{x}_m - \bar{x}_f) + [(\beta_m - \beta^*)\bar{x}_m + (\beta^* - \beta_f)\bar{x}_f] \tag{3}$$

This decomposition can be reduced to Oaxaca's two special cases if it is assumed that there is no discrimination in the male wage structure, i.e. $\beta^* = \beta_m$, or if it is assumed that $\beta^* = \beta_f$. Neumark shows that β^* can be estimated using the weighted average of the wage structures of males and females and advocates using the pooled sample to estimate β^*. The first term is the gender wage gap attributable to differences in characteristics. The second and the third terms capture the difference between the actual and pooled returns for men and women, respectively.

While Neumark's decomposition is attractive, it is not immune from common criticisms of decomposition methods in general, namely, the omission or inadequate measures of variables that affect productivity. Also, without evidence that employers care only about the proportion of each type of labour employed, it is not clear that the pooled coefficient is a good estimator of the non-discriminatory wage structure (Appleton *et al.*, 1999). In addition, like other conventional decomposition methods, Neumark's decomposition fails to account for differences in sectoral structures between gender groups.

Appleton et al. (1999)'s sectoral decomposition

This decomposition technique takes into account sectoral structures between genders. Appleton *et al.* (1999) adopt a similar approach to that of Neumark and decompose the gender wage gap into three components. Since this technique is based on Neumark's decomposition, it does not suffer from the index number problem encountered by previous authors who attempted to account for differences in occupational choices (Brown *et al.*, 1980).

Let $\overline{W}_m$ and $\overline{W}_f$ be the means of the natural logs of male and female earnings and $\bar{p}_{mj}$ and $\bar{p}_{fj}$ be the sample proportions of men and women in sector j respectively. Similarly to

Neumark (1988), Appleton *et al.* (1999) assume a sectoral structure that would prevail in the absence of gender differences in the impact of characteristics on sectoral choice ($\overline{p}_j^*$, the proportion of employees in sector j under this common structure). They then decompose the difference in proportions employed in three sectors such as:

$$\overline{W}_m - \overline{W}_f = \sum_{j=1}^{3} \overline{p}_j^* (\overline{W}_{mj} - \overline{W}_{fj}) + \sum_{j=1}^{3} \overline{W}_{mj} (\overline{p}_{mj} - \overline{p}_j^*) + \sum_{j=1}^{3} \overline{W}_{fj} (\overline{p}_j^* - \overline{p}_{fj}) \qquad (4)$$

A multinominal logit model is used to specify the selection process of an individual into the different sectors. If q_i is a vector of i's relevant characteristics, the probability of a worker i being in sector j is given by:

$$P_{ij} = \exp(\gamma_{ij} q_i) / \sum_{j=1}^{3} \exp(\gamma_{ij} q_i) \text{ with } i = m, f$$

If the distribution of men and women across sectors is determined by the same set of coefficients γ_j^*, then the probability of a worker with characteristics q_i being in sector j is:

$$P_{ij}^* = \exp(\gamma_j^* q_i) / \sum_{j=1}^{3} \exp(\gamma_j^* q_i)$$

Hence, by estimating pooled and separate multinominal logit models for men and women, it is possible to derive the average probability for male and female workers in the different sectors. These mean probabilities are denoted by $\overline{p}_{ij}^*$. The relationship between γ^* and γ_i are similar to that of β^* and β_j in Neumark's decomposition. Embedding the self-selection process in (4), the full decomposition can be written in the following way:

$$\begin{aligned}\overline{W}_m - \overline{W}_f = & \sum_{j=1}^{3} \overline{p}_j^* (\overline{x}_{mj} - \overline{x}_{fj}) \beta_j + \sum_{j=1}^{3} \overline{p}_j^* \overline{x}_{mj} (\beta_{mj} - \beta_j) + \sum_{j=1}^{3} \overline{p}_j^* \overline{x}_{fj} (\beta_j - \beta_{fj}) \\ & + \sum_{j=1}^{3} \overline{W}_{mj} (\overline{p}_{mj}^* - \overline{p}_j^*) + \sum_{j=1}^{3} \overline{W}_{fj} (\overline{p}_j^* - \overline{p}_{fj}^*) + \sum_{j=1}^{3} \overline{W}_{mj} (\overline{p}_{mj} - \overline{p}_{mj}^*) + \sum_{j=1}^{3} \overline{W}_{fj} (\overline{p}_{fj}^* - \overline{p}_{fj}).\end{aligned} \qquad (5)$$

The first three terms are similar to Neumark decompositions of within-sector wage gaps. The fourth and fifth terms measure the difference in earnings due to differences in distribution of male and female workers in different sectors. The last two terms account for differences in earnings resulting from the deviations between predicted and actual sectoral compositions of men and women not accounted for by differences in characteristics.

4. Econometric results

4.1 Potential versus actual experience: refining labour market attachment measures in earnings functions

Tables 2 and 3 in Appendix present the OLS regression estimates of the determinants of log hourly earnings for males and females. A test of equality for the coefficients of male and female earnings functions shows that we have to reject the null hypothesis (at the 1% level) that one single equation can explain both earnings.[15] In column 1, for comparative purposes, we use the individual age as a proxy for the total experience since many studies base their estimates on this variable, for lack of more relevant information. From columns (2) to (10), we then progressively improve this rough measure by replacing the potential measures of experience (columns 1, 2, 4 and 7) by actual ones (columns 3, 5, 6, 8, 9 and 10). Overall, the earnings functions explain, respectively for males and females, about 31% and 52% of the earnings differentials. Note that quadratic and more flexible polynomial specifications for experience and education have been tried but cannot be accurately estimated with this data.[16]

Regressions introducing actual experience instead of potential experience shed light on gender-differentiated effects. From models (2) and (3), when total actual experience is accounted for, the return to experience diminishes for women while it increases for men, and becomes more significant. This is at odds with former expectations. In columns (4) and (5), the same specifications are corrected for potential selectivity bias, employing the method described in section 3.1.[17] In both models, the coefficient on the correction term (IMR) is negative and insignificant at the usual confidence interval (10% level). In other words, the mechanism of allocation in the two groups (paid work participants versus non-participants) does not affect earnings significantly. Finally, the estimated marginal returns to experience are quite small and remain significantly higher for females than for males whatever the estimated model (in column 5, 1.5% against 0.8%). The latter result is a common one, especially in developing countries, given that women generally have less experience than men and are therefore better rewarded for this.

---

[15] First, we ran regressions on the pooled sample (not presented, available upon request) and then performed the Chow test of equality for the male and female coefficients.

[16] We therefore decided to include a quadratic term only when it was sufficiently significant (at the 20% level or better) or for comparative purposes in order to preserve on degrees of freedom and the significance of the other estimated coefficients of experience.

Columns (6) expand the regressors of column (5) to include two labour force attachment variables (LFAVs) reflecting non-working time (total years of unemployment and inactivity). Adding non-working time allows for the possibility that human capital appreciates and depreciates at different rates. Interestingly, neither men nor women seem to be penalised for the time spent as unemployed. However, inactivity is statistically significant in the female regression, but insignificant in the male one. Moreover, quite curiously, females show a positive premium for their periods of inactivity. This is at odds with former intuition but may be explained by socio-economic stylised facts of the Madagascan labour market and/or data deficiencies. Given the confidence we place in the quality of our data, our tentative explanation is that women's inactivity spells are not penalised by employers because the latter may give more value to women's home activity than to unemployment periods strictly speaking. In fact, unemployment is less likely to be related to parenting than inactivity and may voice a negative signal in employers' eyes. In contrast, during women's complete withdrawals from the labour market, there might be a human capital accumulation effect as a result of, for instance, childcare that provides them with parenting skills and more responsibilities in the household. As a result, women returning to work after an absence from the labour market may not necessarily suffer skill losses, nor missed promotion opportunities, compared to their male counterparts who are more likely to work in highly skilled fields where both career advancement and skill depreciation are relatively fast. On the contrary, women may benefit from enhanced credibility. In fact, women's unobserved individual heterogeneities may be positively correlated to their inactivity but also to their earnings.

Introducing inactivity and unemployment periods in earnings functions raises the estimated return to actual experience by 15% for females (from 1.50% to 1.72%) but slightly diminishes that of males (columns 6). We now find that the female return to actual experience is higher than that of potential experience (1.72% versus 1.51%), as previously expected, as was already the case for males. Therefore, given the high amount of time spent out of the labour force for women (on average, 10 years versus 5 years for men), being able to differentiate in earnings functions between the various episodes spent in and out of the labour market seems to be an important step towards refining the returns to human capital variables across the sexes. This may also affect the portion of the gender earnings gap component that is not explained by gender differences in observed characteristics.

[17] The first-stage probit estimates of males and females' employment participation appear in a separate Appendix, not for publication.

We are also interested in the coefficient on the schooling variable. This coefficient, commonly interpreted as the private return to education, seems to be underestimated for females when actual experience is used without controls for limited LFAVs, but remains relatively constant for males (columns 5 and 6). With regard to measurement error, there is in general no reason to suppose that the differential effect of work experience affects estimates of male and female schooling differently. Nonetheless, in the case of work experience, as we showed in sections 1 and 2, the female measure is likely to be subject to relatively large conceptual measurement error. Accordingly, the female experience coefficient may be subject to a relatively large downwards bias. If there is a negative correlation between schooling and work experience (as it is the case in our data), a relatively large downwards bias in the female experience coefficient (as evidenced by our estimates when selectivity and LFAVs are not accounted for, column 3) could in turn give rise to a relatively large downwards bias in the female schooling coefficient (Dougherty, 2003).[18] Columns 2 to 6 of Table 3 highlight that the marginal return to experience was somewhat underestimated for females. Moreover, the estimated coefficients on education are affected by the inclusion of LFAVs, especially for women. Indeed, the return to education increases from 11.10% to 11.96% per year for women, and falls slightly from 8.84% to 8.54% for men. As a consequence, by introducing both actual experience and LFAVs we are able to estimate the "true" return to schooling that may frequently be biased when proxy measures of experience are used in Mincer-type earnings regressions. In particular, the female coefficient on education is likely to be biased downwards when only actual experience is included without controls for other LFAVs. We will use these regressors in the rest of the paper as they have a real impact on the precision of our estimates.

From columns (7) to (10), our purpose is to continue refining the measure of experience by decomposing it into different quantitative and qualitative work spells using individuals' employment records. Columns (7) and (8) take into account the years of tenure with the current employment and its squared value in two types of specification, that is, with potential experience (columns 7) and with actual experience and the two LFAVs (columns 8). In fact, as tenure is an important productivity component and females often have less tenure,

[18] In the model $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u$, where X_1 is subject to measurement error with expected value 0 and variance $\sigma^2_{X_1}$, it can be shown that the limiting value of the OLS estimator of β_2 is:

$$\operatorname{plim}\hat{\beta}_2 = \beta_2 + \frac{\beta_1 \sigma^2_{X_1} \sigma_{X_1X_2}}{\sigma^2_{X_1}\sigma^2_{X_2} - (\sigma_{X_1X_2})^2}$$

where $\sigma^2_{X_1}$ and $\sigma^2_{X_2}$ are the population variances of X_1 and X_2 and $\sigma_{X_1X_2}$ is their population covariance. Given that schooling and work experience tend to be two of the most important variables in wage equations, this relationship may be a guide to the behaviour of the schooling coefficient, despite the multiplicity of additional variables. If $\sigma_{X_1X_2}$ is negative, the bias will be downwards. Note that if work experience neglecting it in a wage regression could lead to a serious over-estimation of the discrimination component, as evidenced by Weichselbaumer and Winter-Ebmer (2005). Our estimates show that tenure and its square are not statistically significant in either model for men, while they are highly significant for women in both specifications. In the case of men, standard human capital theory would interpret model (8) by arguing that general human capital significantly increases wages (the return to total actual experience is statistically significant and positive) unlike specific human capital (the return to tenure is insignificant). However, note that since tenure and total actual experience are positively correlated (0.45), the latter captures a fair amount of the effect of the former when both variables are introduced as regressors.

Columns (9) propose a further alternative measure of experience replacing the commonly used tenure with the current employer. Firstly, we took into account the fact that workers may have had two different and successive occupations with the same employer. For instance, individuals who have worked for the same employer for their entire life may have started as blue-collar workers and, after some time, become white-collar workers. In this context, it might be a strong hypothesis to assume a unique marginal return to tenure for both occupations, even if they took place in the same firm[19] : there is indeed empirical evidence that the returns to tenure may not be entirely sector or firm-specific but also linked to the human capital diffusion process which is, in turn, closely related to workers' occupational features and choices (Destré and Nordman, 2002; Nordman and Hayward, 2004). To address this issue, we tested a variable taking into account the length of the last occupation taken up with the last employer instead of the overall years of tenure with the same employer. The rewards for an additional year of experience in the same occupation amount to 1.19% for males and 3.4% for females at the sample mean, while they are respectively 0.07% and 2.06% for the overall tenure.[20] Hence, it appears that the returns to experience within the last occupation are much higher than those of the overall tenure. The latter may therefore reflect a "sticky floor" effect.[21]

is subject to greater measurement error for females than for males, the differential in the male-female schooling coefficients will be underestimated.

[19] Regarding the gender issue, males have on average more tenure than females (9.2 versus 8.1 years) but both display the same average amount of time spent in the last occupation (about 7.5 years).

[20] Estimates not presented due to lack of space.

[21] At the bottom of the wage distribution in industrialised countries, some authors found that the gender pay gap widens significantly and defined this phenomenon as a sticky floor (see Booth, Francesconi and Frank, 2003). Our data may suggest the existence of this phenomenon in Madagascar. Indeed, lower returns to the total tenure might reflect the fact that a long period of time is needed for the individuals at the bottom of the pay scale to be promoted and to benefit from increased wages.

Secondly, it is debatable whether experience accumulated in the current employment should always be distinguished from that accumulated in previous jobs. Workers may have practiced exactly the same profession, or carried out the same specific duties, in other contexts, firms or workshops (see note 7). Therefore, it could be that with earnings, especially across gender, it is just the time spent in accumulating the technical know-how that is part of each worker's profession that is important and not necessarily where that knowledge was gained.[22] To test this second assumption, we introduced a variable taking into account the time accumulated while working in the same profession in columns (9), e.g. practicing the same duty, irrespective of the workplace, firm or employer. Unlike the males' return to tenure in model (8), the marginal return to experience in the main profession is significant for males, though very low when computing it with the quadratic term at the sample mean (0.6%). For females, this return is lower than that of tenure in model (8) (respectively, 1.48% versus 2.06%). Therefore, the gap between the returns to experience across gender is slightly reduced when experience in the main profession is controlled for. This may indicate that men benefit from their more assiduous participation in employment than women who may suffer more human capital depreciation, or find it difficult to acquire skills related to the same given profession, as a result of their less regular labour force attachment.

Finally, models (10) replace the total actual experience measure by a variable net of the time spent in the current job (thereby, the actual experience off the current main employment) in order to avoid accounting for the same spell of experience twice. The overall actual experience is then segmented and well accounted for. We also introduced additional qualitative LFAVs ('augmented LFAVs'), such as the total number of work interruptions, its squared value – to take into account its possible non-linear marginal effect –, and a dummy indicating workers' high proportion of 'relevant' experience, i.e. whether the proportion of preceding actual experience accumulated in the same sector as the current one is equal to or higher than 50%.[23] We also added an interaction term between the number of work interruptions and the schooling variable to allow for possible differentiated effects of labour market withdrawals across educational levels.[24]

[22] This echoes the question of how to differentiate between the various sources of human capital accumulation in wage equations, which may be different from knowing whether it is general or specific. In fact, the nature of human capital may not be exclusively linked to the fact of belonging to a given employer or firm (i.e. to who pays the worker) but also to what he/she has actually learnt – and how – while performing the same specific task.
[23] This is the ratio of the time spent off the current job working in the same institutional sector as the current one (public, private formal or self-employed/informal) to the total actual experience. The dummy is equal to 1 for 29% of the sample (respectively, 24% and 36% of the sub-samples of males and females).

Both experience variables (actual experience off the incumbent job and tenure) are then statistically significant (except the squared value of tenure for males) which reinforces the idea that the models are better specified when total actual experience is properly segmented as compared to models 8. Note that the dummy for a high proportion of previous actual experience in the same sector (relevant experience) is insignificant for both males and females. This might be an important result since some studies emphasize the importance of relevant experience in wage determination as it is often assumed to be a good measure of job-related human capital (see Barron *et al*., 1999).

Other studies have suggested the potential negative impact of work interruptions on earnings patterns (see section 1), without, however, suggesting compelling estimates mostly due to lack of relevant data. Our estimates suggest that work interruptions have no clear impact on males' earnings (except the quadratic term). Interestingly, these interruptions do affect females' earnings significantly. For women, all the three estimated coefficients are significantly different from zero at the 10% level: negative effect of the number of interruptions, positive impact of its squared value and its interaction with education. Hence, the marginal negative effect of a female's work interruption on her earnings is reduced by: (1) the quantity of these interruptions and (2) her level of education. In other words, highly educated women are less penalised than their poorly educated counterparts. Also, the higher the number of interruptions, the lower the marginal negative effect in absolute value.

Finally, it seems important to consider sectoral participation in order to understand the returns to observed characteristics and, in particular, to human capital variables. The estimates of the sectoral dummies' coefficients are large and often statistically significant (the reference being the public sector that always appears to be the most rewarding for women). This is in accordance with the usual persistence of uncompensated earnings differentials across individuals with identical productive characteristics (Goux and Maurin, 1999). The results show that workers with comparable measured characteristics can have very different earnings because they belong to different institutional sectors. So far, we have disregarded the possible endogeneity of these sectoral participation choices in earnings determination for the sake of simplicity. We now turn to our sectoral approach.

---

[24] The thinking behind this is that the higher the education, the higher the penalty for work interruptions. To our

4.2 Sectoral earnings functions

Estimates from earnings equations for men and women in public wage employment, private wage employment and self-employment or informal sector (hereafter simply "informal") are presented in Tables 4 and 5. The earnings equations are corrected for potential selectivity bias, employing the method described in section 3.1, and using the sectoral choice model estimates to calculate the selectivity factors.[25] For women, only the correction term in the informal sector is positive and statistically significant. This means that unobserved characteristics that increase the probability of working in this sector also have a negative effect on earnings. However, male estimates show that there is sample selectivity for each considered sector: the correction term is significant and negative in both the public and private wage sectors while it is significant and positive in the informal sector. Hence, informal sector participation is associated with unobserved characteristics that are positively correlated to earnings differentials, both for men and women.

As expected, the models' explanatory power goes in descending order from public employment, to private employment, then to informal employment, with R^2 decreasing, depending on the specifications considered, from [0.58, 0.73] to [0.38, 0.40] and [0.29, 0.31] respectively for each of the three sectors. This hierarchy is consistent with the predictions of the standard human capital model, as this is better suited to accounting for the heterogeneity of earnings in the public sector where wages are based on a set scale that takes these criteria (education, tenure) explicitly into account. On the other hand, in the informal sector, apart from the probability of greater measurement errors, other factors not taken into account in our equation, such as the amount of capital, are likely to have a significant impact on earnings. Tests for the joint equality of coefficients (Chow test, likelihood ratios) show that both the decomposition by institutional sector and the separate estimates of equations by gender are justified.

It is in the formal private sector that experience has the most value. Depending on the models, the coefficients of experience vary from 0.0101 to 0.0157 in the public sector and from 0.0229 to 0.0268 in the formal private sector. In the informal sector, on the contrary, actual or potential experience has no significant impact on men's earnings differentials while the returns to women's actual experience amount to 1.3%.

---

knowledge, however, there is no clear theoretical argument to support this intuitive idea.

[25] The maximum likelihood estimates of the multinomial logit sectoral choice models are presented in the separate Appendix, not for publication.

For men and women alike, taking into account the actual number of years worked always leads to an increase in the return to experience, except in the informal sector for men where returns to experience are insignificant. The effect is particularly important for women. For example, in the public sector, one year of additional actual experience leads to an increase in earnings of 1.57%, compared with 1.0% for potential experience.

Education is always a profitable investment and returns are much higher than for experience. Once again, the informal sector is an exception to the rule for men, whose earnings do not depend on their level of schooling. Although non-negligible and weakly significant, the return to education for women is very much lower when actual experience is accounted for, at around 6%, than that recorded for the other sectors. On average, women's education is given more value than that of their male counterparts. This difference reaches 6 percentage points in the informal sector, more than 2 points in the formal private sector and 2 points in the public sector (even 3 points when actual experience and LFAVs are included instead of potential experience). The latter result is all the more surprising that in this sector wages are supposed to follow the same scale for everyone, irrespective of gender. The gap in returns to education may, in part, reflect the impact of occupational segregation or of unobservable factors playing in women's favour. We will take a closer look at this hypothesis in the next section.

Finally, spells of inactivity or unemployment do not seem to penalise workers, except the years of unemployment for males in the public sector, which could indicate labour market segmentation. Whereas, in line with our initial hypothesis, the coefficients are negative for males, i.e. the length of unemployment or withdrawal from the labour market tends to depreciate the human capital of occupied wage-earners, they are often insignificant. On the contrary, and in line with our tentative explanation when commenting the global cross-sector models, women seem to benefit from spells of inactivity in the formal private sector. Having a long-term contract or having received on-the-job training are factors that improve men's earnings, especially in the public sector. Finally, unions have a positive impact on wages, but only in the public sector, the only sector where they have sufficient weight to have effective bargaining power.

4.3 Earnings decompositions

Table 6 provides an overview of the gender earnings decompositions using the alternative decomposition techniques described in section 3.2. We present the main rough results drawn from the decompositions (that is, the proportion of the explained versus unexplained gender earnings gap, hereafter simply "gender gap") using different non-selectivity corrected earnings models. They include our different measures of experience alternatively, the limited and augmented LFAVs together with two sets of *limited* and *augmented* control variables (see definitions at the bottom of Table 6).

The overall results confirm that a greater portion of the gender gap can be explained using actual rather than potential experience. Depending on the decomposition techniques used, the explained component ranges from about 11.4% to 22.5% in the conventional model (using potential experience and the limited control variables) and from 24% to 38.7% using actual experience instead of potential. This variation is quite considerable. Moreover, using the different augmented models discussed in section 4.1 (adding non-working time – the limited and augmented LFAVs – and using successively total actual experience, previous experience off the current job, tenure in the current employment, and experience in the main profession) progressively reduces the share of the unexplained component from 88% to 70.2% in Oaxaca's decomposition. These findings are novel for Madagascar, and more generally for Africa, but similar to findings in developed countries.[26] Hence, the share of the gap attributable to differences in experience between men and women appears to be severely underestimated when potential instead of actual experience variables are used. Looking again at Oaxaca's decomposition, differences in actual experience account for about 9.4% of the gap, while potential work experience explains only 3.3%.

This may be explained as follows: first, as stated earlier, men and women differ little in the mean characteristics of potential experience but they differ significantly in actual experience. Second, although potential and actual experience are highly correlated (0.78), an additional year of actual experience gives different returns than a year of potential experience. So, when the actual measure is used, both the difference in means and the difference in returns produce a greater explained component than when the potential variable is introduced.

Adding time spent out of the labour market (inactivity) and unemployment spells generally increases the percentage of the earnings gap explained by labour market attachment differences (actual experience, inactivity and unemployment). Overall, educational differences continue to explain more of the gender gap than labour force attachment differences (in Oaxaca's decomposition, 22% versus 14%). Interestingly, once actual experience and LFAVs are controlled for, the fraction of the gender gap explained by education remains quite stable. This is at odds with some findings in industrialised countries where, in the absence of actual experience and non-working time spells, Antecol and Bedard (2004) have shown that educational differences appear to absorb some of the systematic differences in labour force attachment.[27] This would suggest that, in the absence of actual experience measures, education is not able to absorb the variations in actual experience since the latter are not necessarily correlated with educational attainment.

Overall, the addition of the various actual experience and non-working time measures increases the proportion of the gender gap explained by observable characteristics to nearly 30% using Oaxaca's decomposition and up to 45% with Neumark's, while using only potential experience allows us to explain no more than 11% and 22% respectively. Hence, Neumark's decomposition clearly always produces the highest share of the explained component.

In the last panel of Table 6, we use the set of augmented control variables including job characteristics such as the type of work contract, the presence of a union, 7 occupational dummies and 9 industry dummies describing the type of activity in the sectors of employment. Unsurprisingly, controlling for these job characteristics greatly reduces the unexplained component of the gender gap. In this context, the unexplained gender gap falls to 61% with Oaxaca's decomposition and to 29% with Neumark's. We would argue that the effect of controlling for job characteristics on the gender earnings gap reflects the occupational segregation that may be present in Madagascar.

Although the selectivity corrections are not significant in the pooled earning equation models, they are sometimes quantitatively large; they also modify the OLS estimates and, hence, may affect the decompositions. Neuman and Oaxaca (2004) show that sample selection complicates the interpretation of wage decompositions. They offer several alternative decompositions, each based on different assumptions and objectives. We use one that consists in considering selectivity as a separate component. This technique has the advantage of not calling for any prior hypothesis regarding the links between individual characteristics and selectivity. An additional term in the decomposition measures the contribution of selection

[26] For example, Wright and Ermisch (1991), O'Neill and Polachek (1993), Myck and Paull (2004) for the United Kingdom and the United States, and Meurs and Ponthieux (2000) for France.
[27] These patterns are drawn from wage gaps across ethnic minorities and race.

effects to the observed gender earnings gap: $\hat{\theta}_m\hat{\lambda}_m - \hat{\theta}_f\hat{\lambda}_f$ where $\hat{\lambda}$ and $\hat{\theta}$ denote respectively the mean IMR and its estimated coefficient from each regression by sex.[28]

Table 7 presents the Oaxaca and Neumark's decompositions controlling for selectivity effects and using two types of earnings model with the limited control variables: potential experience and segmented actual experience plus augmented LFAVs. By means of potential experience, the share of the gap explained by individual characteristics goes from 10% in Oaxaca's decomposition to 19% in Neumark's, while the unexplained portion of the gap amounts respectively to 65% and 56%. In both decompositions, the selectivity component represents 25% of the gender gap. Hence, as compared to our results of Table 6, panel 6, it appears that the selectivity correction has mostly reduced the share of the gap attributed to individual characteristics (respectively, 29% and 45% in Table 6 versus 10% and 19%) instead of diminishing that due to discrimination (respectively, 70% and 54% in Table 6 versus 65% and 56%). Replacing potential experience by actual experience and augmented LFAVs also provides meaningful results: in both decompositions, the portion of the gap explained by individual endowments significantly increases (respectively, to 28% and 33%) to the detriment of the share explained by selectivity effects, which falls to 10%. The discrimination component also diminishes to 62% in Oaxaca's decomposition.

The Neumark decomposition can be used to determine whether the differences in returns reflect higher return for men compared to a pooled (assumed non-discriminatory) structure or lower returns to women. The deviation in female returns from the pooled earnings regression is about ten times more important than the deviation in male returns. Therefore, it can be concluded that discrimination against women is more relevant than nepotism towards men (in Neumark's terminology) in explaining the gender gap.

Of course, these decompositions may suffer from biases due to the sample pooling of individuals working in the three institutional sectors. As we have shown in the previous section, the returns to individual characteristics and especially to human capital variables differ across sectors. Moreover, mean earnings differ greatly across sectors and sexes. This explains why sectoral decompositions induce significant variations in the decompositions across the three sectors.[29] Firstly, in the public sector, mean earnings are higher for women

[28] If the pooled wage structure is used (Neumark, 1988), the selectivity term can be expanded to $\hat{\theta}(\hat{\lambda}_m - \hat{\lambda}_f) + (\hat{\theta}_m - \hat{\theta})\hat{\lambda}_m + (\hat{\theta} - \hat{\theta}_f)\hat{\lambda}_f$, where $\hat{\theta}$ is the estimated IMR coefficient from the pooled sample.

[29] These results are not reproduced to save space but are available upon request.

than for men. In this sector, the gender gap is therefore in favour of females.[30] In fact, women employees have more favourable characteristics than their male counterparts. On the contrary, the gender earnings gap is in favour of males in the private sector and, more importantly, in the informal sectors. In these sectors, respectively 20% and 4% of the gender gap is explained by differences in observed endowments (in favour of males) while workers' characteristics explain much more of the gender gap in the public sector (46%).

Secondly, the same picture as above emerges from Neumark's decomposition for the public and private sectors: from the male and female deviations in returns, discrimination against women is more pronounced than nepotism towards men, especially in the public sector. Given the higher mean earnings for women in this sector, females offset this discrimination in returns by their more favourable observed characteristics. In the informal sector, the deviation in returns from the pooled wage structure is detrimental to men.

However, selectivity effects account for much of the gaps in the public and informal sectors while discrimination is more acute in the private sector. In the informal sector, for instance, selectivity explains almost 90% of the gender gap (Oaxaca's decomposition) while the share attributable to discrimination amounts to 7%.

The full decomposition developed by Appleton *et al.* (1999), taking into account the location of men and women in the three sectors, is finally presented in Table 8. We control for selectivity effects using earnings offered to men and women (instead of actual earnings) which are net of the impact of the selectivity corrections, that is, $(\overline{W}_{mj} - \hat{\theta}_{mj}\hat{\lambda}_{mj})$ and $(\overline{W}_{fj} - \hat{\theta}_{fj}\hat{\lambda}_{fj})$ for the *j* sectors (see Reimer, 1983; Appleton *et al.*, 1999). The first three terms address the differences in returns due to within-sector differences and are weighted sums of the Neumark's decomposition of the within-sector earnings gaps. In line with the traditional decomposition results of Table 7, the deviations in returns (the discrimination component) explain much of the within-sector differences. The same picture emerges from Appleton *et al.*'s full decomposition on Côte d'Ivoire, which also show negative signs on the deviation components, that is to say, favourable deviation of females' returns as compared to the pooled earnings structure.

[30] Similarly, Glewwe (1990) found no wage discrimination against women in the public sector in Ghana.

The last three terms of the full decomposition tell us the share of the gender gap which may be attributed to gender differences in proportions of workers in each sector. The positive sum of these three terms implies that the differences in sectoral locations are more favourable to men than to women. The gender earnings gap would have been more than three times smaller if men and women had been equally distributed across the three sectors. This might be because fewer women than men are located in the higher paying public sector where the gender earnings gap is in favour of women. Female paid work participants are found less in the public sector than their male counterparts (respectively, 35% against 64%) while they are almost equally distributed in the lower paying informal sector (49% versus 51%). Hence, the weak representation of women in the higher paying public sector appears to contribute towards keeping the gender pay gap greater than it otherwise would be.

5. Conclusion

Our study of Madagascar represents the first attempt to shed light on the determination of the gender earnings gap while using detailed information from biographical and labour force surveys. This unique matched data set enables us to reassess the returns to human capital across gender, notably by introducing various measures of individuals' labour force attachment. We then propose different decompositions of the gender earnings gap that take into account (1) the effects of selection relating to labour force and sectoral participations (public, formal private and informal/self-employed sectors) and (2) alternatives to the standard methods for measuring human capital, especially workers' professional experience.

Our results show that, although the experience coefficients from earnings regressions based on potential and actual experience are almost similar when these variables are introduced alone, adding more detailed labour force attachment variables (unemployment, inactivity spells or the number of work interruptions) leads us to greatly reassess these estimates. Using these regressors in earnings functions increases the return to actual experience for both males and females. This return always exceeds that of potential experience. In addition, we found a negative effect of the number of work interruptions on females' earnings. This marginal negative effect decreases with the quantity of interruptions. Also, with regard to labour force withdrawals, highly educated women seem less penalised than their poorly educated counterparts.

The estimates segmented by sector also highlight that, for men and women alike, taking into account the actual number of years worked always leads to an increase in the return to experience. The effect is particularly important for women. For example, in the public sector, one year of actual additional experience leads to an increase in earnings of 1.57%, compared with 1.0% for potential experience. Spells of inactivity or unemployment do not seem to penalise workers, except the years of unemployment for males in the public sector.

Our various earnings decompositions show that differences in average actual experience across sexes lead to markedly different estimates of the fraction of the gender earnings gap that is explained by experience. In non-selectivity corrected earnings decompositions, the addition of the different actual experience and non-working time measures increases the proportion of the gender gap explained by observable characteristics to nearly 30% using Oaxaca's decomposition and up to 45% with Neumark's decomposition, while using only potential experience allows us to explain no more than 11% and 22% respectively. We also provide evidence that, in the absence of labour force attachment measures, education is not able to absorb the variations in actual experience since the latter are not necessarily correlated with educational attainment. This is an additional argument to support the need for more precise labour force participation measures in developing countries. Once sample selectivity effects are controlled for, replacing potential experience by actual experience and labour force attachment variables still provides meaningful results: in both Oaxaca and Neumark's decompositions, the portion of the gap explained by individual endowments increases significantly to the detriment of the share explained by selectivity effects.

The gender earnings decomposition also differs across sectors. The gender gap is in favour of males in the formal private and informal sectors while, in the public sector, women seem better off than men. Yet, selectivity effects account for much of the gaps in the public and informal sectors while discrimination is more acute in the formal private sector. However, traditional decomposition methods fail to account for differences in sectoral structures between gender groups. We therefore utilise Appleton *et al.* (1999)'s decomposition technique which incorporates the impact of sectoral location to examine the gender earnings disparities within each sector. They reveal that the differences in sectoral locations are more favourable to men than to women. The gender earnings gap would have been more than three times smaller if men and women had been equally distributed across the three sectors. Hence, the weak representation of women in the higher paying public sector appears to contribute towards keeping the gender pay gap greater than it otherwise would be. Therefore, public sector downsizing (the partial freeze on public sector recruitment from the mid-1980s in

Madagascar) worsens women's economic position as more women move away from the state sector to the private sector. The separate decompositions by sector, such as Oaxaca and Neumark's, ignore sectoral composition differences, masking the extent of the impact of the state sector downsizing on women.

Nonetheless, in spite of the refinements of the labour force attachment measures we suggested, the regression models used in the decomposition analysis account for no more than half of the variation in the earnings of men and women. The model might be better fitted to the data by including other variables deemed to influence earnings. Typically, the data used comes from household surveys. For a long time, researchers have been unable to document the potential effect of job and firm characteristics – other than industry and firm size – on the wages of men and women. New linked employer–employee surveys would therefore allow researchers to move beyond the individual worker to consider the importance of the workplace in wage determination. There is much to learn about the demand-side factors that may influence employers when they make decisions concerning hiring and promotions or use gender to predict future work commitment. There is clearly still room for prolific studies in this direction.

Table 1. Main descriptive statistics for paid work participants

Variables	**Males** (n=1 063)		**Females** (n=827)	
	Mean	Std. Dev.	Mean	Std. Dev.
Average age	40.28	8.17	40.24	8.26
Average schooling successfully completed	8.87	4.38	7.89	4.35
Average schooling (time spent in school)	11.69	5.87	10.23	5.73
Potential work experience (age – schooling – 6)	22.60	10.24	24.01	10.72
Actual labour market experience	20.58	9.70	17.18	10.51
Actual labour market experience off the current job	11.52	9.20	9.62	9.18
Tenure with the current employer	9.24	8.43	8.08	8.05
Unemployment periods	1.14	2.18	0.82	1.90
Inactivity periods	5.52	4.12	10.84	9.44
Number of work interruptions	0.73	0.97	0.73	0.83
Proportion of previous experience in the same sector exceeding 50% (1 if yes; 0 otherwise)	0.24	0.43	0.36	0.48
Catholic (1 if yes; 0 otherwise)	0.39	0.49	0.38	0.49
Merina (1 if yes; 0 otherwise)	0.88	0.32	0.90	0.30
Married (1 if yes; 0 otherwise)	0.82	0.38	0.64	0.48
Formal training received in the current job (1 ; 0 otherwise)	0.14	0.35	0.10	0.29
Number of hours worked per week	45.44	16.84	40.34	17.44
Short-term contract (CDD) (1 if yes; 0 otherwise)	0.08	0.28	0.04	0.20
Long-term contract (CDI) (1 if yes; 0 otherwise)	0.34	0.47	0.29	0.45
Presence of union in the current job (1 if yes; 0 otherwise)	0.21	0.41	0.17	0.37
Public employment (1 if yes; 0 otherwise)	0.25	0.44	0.18	0.38
Formal private employment (1 if yes; 0 otherwise)	0.30	0.46	0.28	0.45
Self-employment or informal sector (1 if yes; 0 otherwise)	0.44	0.50	0.54	0.50

Sources: *Enquête 1-2-3*, *Phase 1*, *1998*, *Biomad98*, MADIO; authors' calculations.

Table 2. Earnings Functions for Males

Dependent variable: Log hourly earnings

	OLS	**OLS**	**OLS**	**OLS with Selectivity correction**	**OLS with Selectivity correction**	**OLS with Selectivity correction**	**OLS with Selectivity correction**	**OLS with Selectivity correction**	**OLS with Selectivity correction**	**OLS with Selectivity correction**
	Age as potential experience	*Potential experience*	*Actual experience*	*Potential experience*	*Actual experience*	*Actual experience + limited LFAV*	*Potential experience + tenure*	*Actual experience + tenure + limited LFAV*	*Actual experience + 'profession' + limited LFAV*	*Actual experiences + augmented LFAV*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Years of completed schooling	0.0801***	0.0878***	0.0886***	0.0880***	0.0884***	0.0854***	0.0875***	0.0851***	0.0852***	0.0912***
	(13.95)	(12.84)	(13.54)	(12.17)	(12.77)	(11.82)	(12.15)	(11.82)	(11.83)	(11.28)
Age	0.0083***	–	–	–	–	–	–	–	–	–
	(2.68)									
Potential years of labour market experience	–	0.0071**	–	0.0067**	–	–	0.0046	–	–	–
		(2.50)		(2.20)			(1.40)			
Actual years of labour market experience	–	–	0.0091***	–	0.0086***	0.0081***	–	0.0062*	0.0071**	–
			(3.16)		(2.81)	(2.58)		(1.85)	(2.08)	
Actual years of experience off the current employment	–	–	–	–	–	–	–	–	–	0.0057
										(1.60)
Years of tenure with the current employment	–	–	–	–	–	–	0.0111	0.0114	–	0.0194**
							(1.25)	(1.29)		(1.98)
(Years of tenure with the current employment)^2/100	–	–	–	–	–	–	-0.0190	-0.0228	–	-0.0284
							(0.64)	(0.76)		(0.92)
Years of experience in the main profession	–	–	–	–	–	–	–	–	0.0162*	–
									(1.85)	
(Years of experience in the main profession)^2/100	–	–	–	–	–	–	–	–	-0.0461	–
									(1.56)	
Unemployment spells (in years)	–	–	–	–	–	-0.0122	–	-0.0127	-0.0131	-0.0114
						(1.18)		(1.21)	(1.26)	(1.05)
Total inactivity spells apart from unemployment (in years)	–	–	–	–	–	-0.0059	–	-0.0065	-0.0063	-0.0063
						(0.92)		(1.01)	(0.99)	(0.98)
Total number of work interruptions	–	–	–	–	–	–	–	–	–	-0.0405
										(0.54)
(Total number of work	–	–	–	–	–	–	–	–	–	0.0268*

interruptions)^2										(1.84)
Total number of work interruptions × years of completed schooling	–	–	–	–	–	–	–	–	–	-0.0061
										(1.11)
Proportion of previous experience in the same sector exceeding50% (1 if yes; 0 otherwise)	–	–	–	–	–	–	–	–	–	0.0420
										(0.66)
Catholic (1 if yes; 0 otherwise)	-0.0834*	-0.0849*	-0.0851*	-0.0878*	-0.0850*	-0.0828*	-0.0796	-0.0755	-0.0774	-0.0698
	(1.74)	(1.77)	(1.77)	(1.75)	(1.69)	(1.65)	(1.58)	(1.50)	(1.55)	(1.38)
Merina (1 if yes; 0 otherwise)	-0.1372*	-0.1393*	-0.1410*	-0.1721**	-0.1714**	-0.1714**	-0.1719**	-0.1721**	-0.1730**	-0.1755**
	(1.87)	(1.89)	(1.92)	(2.05)	(2.03)	(2.02)	(2.04)	(2.02)	(2.03)	(2.06)
Married (1 if yes; 0 otherwise)	0.2446***	0.2468***	0.2376***	0.2089***	0.2138***	0.2242***	0.2095***	0.2234***	0.2196***	0.2293***
	(3.91)	(3.95)	(3.81)	(3.01)	(3.10)	(3.28)	(3.03)	(3.27)	(3.21)	(3.40)
Formal training received in the current job (1 if received; 0 otherwise)	0.1705**	0.1712**	0.1722**	0.1676***	0.1691***	0.1660***	0.1513**	0.1508**	0.1741***	0.1516**
	(2.36)	(2.37)	(2.38)	(2.76)	(2.80)	(2.73)	(2.52)	(2.50)	(2.85)	(2.53)
Number of hours worked per week	-0.0189***	-0.0190***	-0.0189***	-0.0194***	-0.0193***	-0.0193***	-0.0194***	-0.0193***	-0.0193***	-0.0194***
	(13.45)	(13.59)	(13.55)	(11.93)	(11.88)	(11.88)	(11.88)	(11.84)	(11.84)	(11.86)
Short-term contract (CDD) (1 if yes; 0 otherwise)	-0.1166	-0.1216	-0.1138	-0.1210	-0.1146	-0.1222	-0.1060	-0.1093	-0.1050	-0.0923
	(1.18)	(1.23)	(1.16)	(1.43)	(1.35)	(1.44)	(1.26)	(1.30)	(1.23)	(1.10)
Long-term contract (CDI) (1 if yes; 0 otherwise)	0.0651	0.0644	0.0648	0.0669	0.0671	0.0695	0.0677	0.0694	0.0688	0.0747
	(0.90)	(0.89)	(0.90)	(1.06)	(1.07)	(1.11)	(1.08)	(1.11)	(1.10)	(1.18)
Presence of union in the current job (1 if yes; 0 otherwise)	0.1022	0.0987	0.0988	0.0991**	0.0987**	0.1026**	0.0840*	0.0880*	0.0961*	0.0918*
	(1.48)	(1.42)	(1.43)	(2.00)	(2.01)	(2.07)	(1.71)	(1.78)	(1.95)	(1.87)
Formal private wage employment (1 if yes; 0 otherwise)	-0.0706	-0.0760	-0.0678	-0.0621	-0.0560	-0.0625	-0.0357	-0.0382	-0.0517	-0.0252
	(0.99)	(1.07)	(0.95)	(1.13)	(1.03)	(1.14)	(0.62)	(0.67)	(0.94)	(0.44)
Self-employment or informal sector (1 if yes; 0 otherwise)	-0.1871**	-0.1977**	-0.1885**	-0.1877**	-0.1799**	-0.1831**	-0.1603**	-0.1598**	-0.1751**	-0.1476*
	(2.08)	(2.20)	(2.10)	(2.44)	(2.34)	(2.39)	(2.01)	(2.01)	(2.27)	(1.86)
IMR_{males}	–	–	–	-0.1980	-0.1396	-0.1107	-0.1822	-0.1003	-0.0822	-0.0998
				(1.25)	(0.88)	(0.69)	(1.15)	(0.62)	(0.52)	(0.62)
Constant	0.0372	0.1608	0.1266	0.2636	0.2128	0.2864	0.2205	0.2451	0.2183	0.1947
	(0.19)	(0.91)	(0.74)	(1.43)	(1.17)	(1.53)	(1.17)	(1.28)	(1.12)	(1.01)
Observations	1063	1063	1063	1051	1051	1051	1051	1051	1051	1051
R-squared	0.3844	0.3838	0.3860	0.3881	0.3898	0.3910	0.3902	0.3928	0.3931	0.3961

Absolute value of *t* statistics are in brackets. ***, ** and * mean respectively significant at the 1%, 5% and 10% levels.

Table 3. Earnings Functions for Females

Dependent variable: Log hourly earnings

	OLS	OLS	OLS	OLS with Selectivity correction	OLS with Selectivity correction	OLS with Selectivity correction	OLS with Selectivity correction	OLS with Selectivity correction	OLS with Selectivity correction	OLS with Selectivity correction
	Age as potential experience	*Potential experience*	*Actual experience*	*Potential experience*	*Actual experience*	*Actual experience + limited LFAV*	*Potential experience + tenure*	*Actual experience + tenure + limited LFAV*	*Actual experience + 'profession' + limited LFAV*	*Actual experiences + augmented LFAV*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Years of completed schooling	0.1041*** (14.00)	0.1219*** (13.68)	0.1091*** (14.06)	0.1214*** (11.76)	0.1110*** (12.19)	0.1196*** (11.65)	0.1209*** (11.93)	0.1192*** (11.78)	0.1188*** (11.72)	0.0992*** (8.28)
Age	0.0169*** (5.07)	–	–	–	–	–	–	–	–	–
Potential years of labour market experience	–	0.0160*** (5.13)	–	0.0151*** (4.18)	–	–	0.0129*** (3.31)	–	–	–
Actual years of labour market experience	–	–	0.0134*** (4.95)	–	0.0150*** (4.01)	0.0172*** (4.39)	–	0.0145*** (3.33)	0.0163*** (3.71)	–
Actual years of experience off the current employment	–	–	–	–	–	–	–	–	–	0.0119** (2.46)
Years of tenure with the current employment	–	–	–	–	–	–	0.0418*** (4.02)	0.0415*** (3.97)	–	0.0560*** (5.07)
(Years of tenure with the current employment)^2/100	–	–	–	–	–	–	-0.1287*** (3.35)	-0.1290*** (3.29)	–	-0.1375*** (3.49)
Years of experience in the main profession	–	–	–	–	–	–	–	–	0.0398*** (3.92)	–
(Years of experience in the main profession)^2/100	–	–	–	–	–	–	–	–	-0.1337*** (3.70)	–
Unemployment spells (in years)	–	–	–	–	–	0.0036 (0.28)	–	0.0021 (0.16)	0.0036 (0.28)	-0.0036 (0.28)
Total inactivity spells apart from unemployment (in years)	–	–	–	–	–	0.0078** (1.97)	–	0.0069* (1.73)	0.0078** (1.98)	0.0055 (1.40)
Total number of work interruptions	–	–	–	–	–	–	–	–	–	-0.1965* (1.95)
(Total number of work	–	–	–	–	–	–	–	–	–	0.0347*

interruptions)^2										(1.78)
Total number of work interruptions × years of completed schooling	–	–	–	–	–	–	–	–	–	0.0212**
										(2.57)
Proportion of previous experience in the same sector exceeding 50% (1 if yes; 0 otherwise)	–	–	–	–	–	–	–	–	–	0.0018
										(0.03)
Catholic	0.0396	0.0411	0.0380	0.0343	0.0325	0.0317	0.0470	0.0441	0.0304	0.0392
(1 if yes; 0 otherwise)	(0.74)	(0.76)	(0.71)	(0.63)	(0.60)	(0.58)	(0.88)	(0.83)	(0.57)	(0.74)
Merina	-0.1216	-0.1330	-0.1182	-0.1558*	-0.1089	-0.1232	-0.1443	-0.1182	-0.1209	-0.1159
(1 if yes; 0 otherwise)	(1.38)	(1.51)	(1.34)	(1.67)	(1.16)	(1.33)	(1.57)	(1.29)	(1.33)	(1.26)
Married	0.1581***	0.1543***	0.1644***	0.1800***	0.1469**	0.1557***	0.1558***	0.1369**	0.1403**	0.1468***
(1 if yes; 0 otherwise)	(2.90)	(2.83)	(3.01)	(3.27)	(2.57)	(2.77)	(2.88)	(2.48)	(2.52)	(2.69)
Formal training received in the	0.0875	0.0933	0.0671	0.0883	0.0642	0.0764	0.0520	0.0438	0.0744	0.0530
current job (1 if received; 0 otherwise)	(0.90)	(0.96)	(0.69)	(1.49)	(1.11)	(1.30)	(0.90)	(0.76)	(1.33)	(0.93)
Number of hours worked per week	-0.0173***	-0.0173***	-0.0180***	-0.0172***	-0.0178***	-0.0177***	-0.0172***	-0.0176***	-0.0180***	-0.0177***
	(11.40)	(11.43)	(11.83)	(10.31)	(10.62)	(10.52)	(10.42)	(10.56)	(10.82)	(10.50)
Short-term contract (CDD)	-0.0119	-0.0225	-0.0316	-0.0221	-0.0427	-0.0259	0.0059	0.0030	0.0233	0.0056
(1 if yes; 0 otherwise)	(0.08)	(0.16)	(0.22)	(0.17)	(0.32)	(0.20)	(0.04)	(0.02)	(0.18)	(0.04)
Long-term contract (CDI)	-0.0587	-0.0694	-0.0657	-0.0765	-0.0718	-0.0721	-0.0950	-0.0901	-0.0602	-0.0779
(1 if yes; 0 otherwise)	(0.67)	(0.79)	(0.75)	(1.07)	(1.01)	(1.02)	(1.33)	(1.27)	(0.85)	(1.09)
Presence of union in the current job	0.1179	0.1230	0.1115	0.1146**	0.1128**	0.1132**	0.1141**	0.1131**	0.1103**	0.1208**
(1 if yes; 0 otherwise)	(1.35)	(1.41)	(1.27)	(2.12)	(2.09)	(2.11)	(2.14)	(2.12)	(2.06)	(2.25)
Formal private wage employment	-0.1453	-0.1488*	-0.1675*	-0.1346**	-0.1672***	-0.1379**	-0.0767	-0.0811	-0.0991	-0.1019*
(1 if yes; 0 otherwise)	(1.64)	(1.68)	(1.91)	(2.18)	(2.82)	(2.25)	(1.22)	(1.30)	(1.59)	(1.65)
Self-employment or informal	-0.6328***	-0.6497***	-0.6552***	-0.6364***	-0.6586***	-0.6365***	-0.5797***	-0.5810***	-0.5835***	-0.6029***
sector (1 if yes; 0 otherwise)	(5.74)	(5.93)	(5.98)	(6.99)	(7.32)	(7.06)	(6.33)	(6.42)	(6.40)	(6.62)
$IMR_{females}$	–	–	–	-0.1989	0.0682	-0.0161	-0.1290	0.0139	0.0097	-0.0264
				(1.49)	(0.41)	(0.10)	(0.94)	(0.08)	(0.06)	(0.16)
Constant	-0.6291***	-0.4412**	-0.1649	-0.3404	-0.2377	-0.4148*	-0.5195**	-0.5716**	-0.5856**	-0.3558
	(2.65)	(2.08)	(0.88)	(1.52)	(1.12)	(1.78)	(2.32)	(2.45)	(2.51)	(1.48)
Observations	827	827	827	823	823	823	823	823	823	823
R-squared	0.5101	0.5105	0.5095	0.5134	0.5106	0.5133	0.5248	0.5244	0.5243	0.5294

Absolute value of *t* statistics are in brackets. ***, ** and * mean respectively significant at the 1%, 5% and 10% levels.

Table 4. Selectivity Corrected Log Earnings Functions Across Sectors for Males
Dependent variable: Log hourly earnings

	Public wage employment			Formal private wage employment			Self-employed or informal sector		
	Potential experience	*Actual experience*	*Actual + limited LFAV*	*Potential experience*	*Actual experience*	*Actual + limited LFAV*	*Potential experience*	*Actual experience*	*Actual + limited LFAV*
Years of completed schooling	0.0658***	0.0684***	0.0578***	0.1029***	0.0998***	0.0999***	0.0180	0.0243	0.0161
	(4.73)	(5.32)	(4.20)	(8.20)	(7.97)	(7.74)	(0.85)	(1.20)	(0.75)
Potential years of experience	0.0101*	–	–	0.0229***	–	–	-0.0069	–	–
	(1.88)			(4.70)			(1.41)		
Actual years of experience	–	0.0132***	0.0114**	–	0.0241***	0.0237***	–	-0.0039	-0.0057
		(2.71)	(2.23)		(4.65)	(4.45)		(0.81)	(1.13)
Unemployment years	–	–	-0.0249*	–	–	-0.0088	–	–	-0.0238
(in years)			(1.77)			(0.43)			(1.63)
Total inactivity years apart from	–	–	-0.0130	–	–	0.0017	–	–	-0.0100
unemployment (in years)			(1.49)			(0.17)			(0.84)
Catholic	-0.1286**	-0.1185*	-0.1156*	-0.0998	-0.1149	-0.1175	-0.0594	-0.0634	-0.0598
(1 if yes; 0 otherwise)	(2.08)	(1.92)	(1.88)	(1.24)	(1.44)	(1.46)	(0.65)	(0.69)	(0.65)
Merina	-0.0270	-0.0347	-0.0290	-0.2110	-0.2179	-0.2125	-0.1423	-0.1519	-0.1338
(1 if yes; 0 otherwise)	(0.31)	(0.39)	(0.32)	(1.43)	(1.48)	(1.42)	(0.81)	(0.87)	(0.75)
Married	-0.0027	-0.0008	0.0025	0.0411	0.0509	0.0475	0.2225**	0.2195**	0.2098*
(1 if yes; 0 otherwise)	(0.03)	(0.01)	(0.02)	(0.38)	(0.48)	(0.44)	(2.08)	(2.05)	(1.94)
Formal training received in the current job	0.2238***	0.2282***	0.2154***	0.0623	0.0780	0.0781	0.3333*	0.3446*	0.3242*
(1 if received; 0 otherwise)	(3.58)	(3.62)	(3.45)	(0.50)	(0.63)	(0.63)	(1.77)	(1.83)	(1.71)
Number of hours worked per week	-0.0231***	-0.0227***	-0.0222***	-0.0213***	-0.0221***	-0.0221***	-0.0173***	-0.0173***	-0.0173***
	(8.68)	(8.60)	(8.51)	(6.17)	(6.24)	(6.24)	(7.97)	(7.96)	(7.97)
Short-term contract (CDD)	-0.0240	-0.0097	-0.0180	-0.0268	-0.0210	-0.0244	-0.0825	-0.0762	-0.1063
(1 if yes; 0 otherwise)	(0.19)	(0.08)	(0.14)	(0.22)	(0.18)	(0.20)	(0.31)	(0.28)	(0.39)
Long-term contract (CDI)	0.2039***	0.2043***	0.2041**	0.0317	0.0336	0.0351	-0.0889	-0.0917	-0.0544
(1 if yes; 0 otherwise)	(2.62)	(2.63)	(2.57)	(0.34)	(0.36)	(0.37)	(0.61)	(0.62)	(0.38)
Presence of union in the current job	0.1250*	0.1227*	0.1303**	-0.0035	0.0051	0.0047	0.2054	0.1460	0.2206
(1 if yes; 0 otherwise)	(1.92)	(1.94)	(2.06)	(0.04)	(0.06)	(0.06)	(1.44)	(1.13)	(1.27)
$\lambda_{m,public}$	-1.2257**	-1.0796**	-1.3276***	–	–	–	–	–	–
	(2.40)	(2.12)	(2.63)						

$\lambda_{m,\,formal}$	–	–	–	-1.4457**	-1.4176**	-1.4160**	–	–	–
				(2.42)	(2.33)	(2.30)			
$\lambda_{m,\,informal}$	–	–	–	–	–	–	2.7155***	2.6216***	2.7475***
							(3.97)	(3.90)	(4.01)
Constant	1.2064**	1.0225**	1.3919***	0.8192	0.8866*	0.8921*	-0.7792**	-0.8420***	-0.7343**
	(2.34)	(2.09)	(2.75)	(1.62)	(1.75)	(1.76)	(2.41)	(2.59)	(2.25)
Observations	270	270	270	321	321	321	460	460	460
R-squared	0.58	0.58	0.59	0.38	0.39	0.39	0.29	0.29	0.30

Robust *t* statistics are in brackets. ***, ** and * mean respectively significant at the 1%, 5% and 10% levels.

Table 5. Selectivity Corrected Log Earnings Functions Across Sectors for Females

Dependent variable: Log hourly earnings

	Public wage employment			Formal private wage employment			Self-employed or informal sector		
	Potential experience	*Actual experience*	*Actual + limited LFAV*	*Potential experience*	*Actual experience*	*Actual + limited LFAV*	*Potential experience*	*Actual experience*	*Actual + limited LFAV*
Years of completed schooling	0.0814***	0.0846***	0.0892***	0.1289***	0.1138***	0.1267***	0.0868***	0.0629***	0.0619**
	(4.77)	(4.84)	(4.96)	(7.08)	(7.03)	(7.06)	(3.51)	(2.86)	(2.35)
Potential years of experience	0.0100*	–	–	0.0265***	–	–	0.0112**	–	–
	(1.96)			(4.96)			(2.13)		
Actual years of experience	–	0.0131**	0.0157***	–	0.0230***	0.0268***	–	0.0135***	0.0130**
		(2.46)	(2.71)		(5.01)	(5.43)		(3.32)	(2.49)
Unemployment years (in years)	–	–	0.0214*	–	–	0.0101	–	–	-0.0047
			(1.88)			(0.36)			(0.29)
Total inactivity years apart from unemployment (in years)	–	–	0.0003	–	–	0.0156**	–	–	-0.0005
			(0.05)			(2.18)			(0.09)
Catholic (1 if yes; 0 otherwise)	-0.0242	-0.0150	-0.0092	0.1582*	0.1301	0.1462	0.0355	0.0546	0.0546
	(0.39)	(0.25)	(0.15)	(1.66)	(1.34)	(1.51)	(0.41)	(0.64)	(0.64)
Merina (1 if yes; 0 otherwise)	0.0855	0.0780	0.0541	-0.0108	-0.0062	-0.0114	-0.2149	-0.1891	-0.1885
	(0.92)	(0.80)	(0.52)	(0.06)	(0.03)	(0.06)	(1.51)	(1.32)	(1.31)
Married (1 if yes; 0 otherwise)	0.1315**	0.1130*	0.1107*	0.1270	0.1588*	0.1294	0.1136	0.1100	0.1084
	(2.31)	(1.96)	(1.92)	(1.52)	(1.91)	(1.55)	(1.41)	(1.36)	(1.36)

Formal training received in the current job (1 if received; 0 otherwise)	0.0681 (1.05)	0.0455 (0.70)	0.0472 (0.73)	0.1320 (1.31)	0.1077 (1.10)	0.1047 (1.06)	0.2252 (1.18)	0.2160 (1.40)	0.2194 (1.37)
Number of hours worked per week	-0.0352*** (8.21)	-0.0356*** (8.28)	-0.0357*** (8.23)	-0.0159*** (3.89)	-0.0161*** (3.90)	-0.0161*** (3.91)	-0.0154*** (8.06)	-0.0161*** (8.40)	-0.0162*** (8.39)
Short-term contract (CDD) (1 if yes; 0 otherwise)	0.1314 (1.01)	0.1488 (1.13)	0.1506 (1.13)	-0.0730 (0.35)	-0.0919 (0.44)	-0.0658 (0.33)	0.6921** (2.30)	0.6426** (2.11)	0.6375** (2.06)
Long-term contract (CDI) (1 if yes; 0 otherwise)	0.1134 (1.26)	0.1287 (1.47)	0.1094 (1.25)	-0.0722 (0.75)	-0.0937 (0.97)	-0.0739 (0.77)	-0.2344 (1.09)	-0.1641 (0.90)	-0.1576 (0.86)
Presence of union in the current job (1 if yes; 0 otherwise)	0.2257*** (3.45)	0.2281*** (3.48)	0.2391*** (3.55)	-0.0356 (0.48)	-0.0618 (0.82)	-0.0493 (0.67)	0.0000 (.)	0.0000 (.)	0.0000 (.)
$\lambda_{m,public}$	-0.2934 (0.79)	-0.0338 (0.08)	0.0832 (0.19)	–	–	–	–	–	–
$\lambda_{m,formal}$	–	–	–	-0.7768 (1.30)	-0.2734 (0.46)	-0.6406 (1.03)	–	–	–
$\lambda_{m,informal}$	–	–	–	–	–	–	1.3980** (2.27)	2.0645*** (3.32)	2.0752*** (3.22)
Constant	0.6366 (1.39)	0.4500 (0.93)	0.2934 (0.58)	-0.5525 (1.00)	-0.5071 (0.91)	-0.6075 (1.09)	-1.4409*** (4.40)	-1.5802*** (5.18)	-1.5599*** (4.64)
Observations	148	148	148	232	232	232	443	443	443
R-squared	0.72	0.73	0.73	0.40	0.39	0.40	0.30	0.31	0.31

Robust *t* statistics are in brackets. ***, ** and * mean respectively significant at the 1%, 5% and 10% levels.

Table 6. Overview of Gender Earnings Decompositions Using Alternative Decomposition Techniques and Non-Selectivity Corrected Earnings Models

Earnings Functions with	**Oaxaca (1973) Blinder (1973)***		**Reimers (1983)**		**Cotton (1988)**		**Neumark (1988) Oaxaca and Ransom (1994)****	
	% unexplained	% explained	% unexplained	% explained	% unexplained	% explained	% unexplained	% explained
Potential experience + *limited* control variables [a]	88.6	11.4	82.8	17.2	83.5	16.5	77.5	22.5
Total actual experience + *limited* control variables	76.1	23.9	69.0	31.0	69.8	30.2	61.3	38.7
Total actual experience + *limited LFAV* + *limited* control variables	71.9	28.1	69.1	30.9	69.5	30.5	56.2	43.8
Segmented actual experience *(experience off the job + tenure)* + *limited LFAV* + *limited* control variables	72.8	27.2	70.0	30.0	70.3	29.7	56.3	43.7
Segmented actual experience *(other experience + experience of the main profession)* + *limited LFAV* + *limited* control variables	72.1	27.9	69.5	30.5	69.8	30.2	56.1	43.9
Segmented actual experience *(experience off the job + tenure)* + *augmented LFAV* + *limited* control variables	70.2	29.8	67.2	32.8	67.5	32.5	54.6	45.4
Segmented actual experience *(experience off the job + tenure)* + *augmented LFAV* + *augmented* control variables [b]	61.1	38.9	47.0	53.0	48.8	51.2	29.5	70.5

Notes: *: the male earnings are taken as the non-discriminatory structure (Ω=1). **: pooled model for both sexes. *a*: this includes education plus all control variables introduced in models 1 to 5 of Tables 2 and 3 minus CDD, CDI, the dummy for union and the sectoral dummies. *b*: this includes the *limited* control variables plus CDD, CDI, a dummy for union, 9 industry and 7 occupational dummies.

Table 7. The Oaxaca and Neumark Decompositions Using Selectivity Corrected Earnings Models

Gender earnings gap decompositions	Earnings model with potential experience	Earnings model with segmented actual experience + augmented LFAV
Oaxaca's decomposition*		
Differences		
Due to characteristics (E)	0.040	0.116
Due to returns to characteristics (C)	-0.806	-0.811
Shift in constant coefficients (U)	1.079	1.069
Selectivity (S)	0.105	0.044
Raw differential (R): E+C+U+S	0.418	0.418
Due to discrimination (D): C+U	0.273	0.258
Effect of selectivity as % total (S/R)	**25.09**	**10.51**
Endowments as % total (E/R)	**9.61**	**27.74**
Discrimination as % total (D/R)	**65.30**	**61.75**
Total	100	100
Neumark's decomposition		
Differences		
Due to characteristics (E)	0.078	0.137
Due to deviation of male returns (C1)	0.015	0.026
Due to deviation of female returns (C2)	0.220	0.211
Selectivity (S)	0.105	0.044
Raw differential (R): E+C1+C2+S	0.418	0.418
Due to discrimination (D): C1+C2	0.235	0.237
Effect of selectivity as % total (S/R)	**25.09**	**10.51**
Endowments as % total (E/R)	**18.60**	**32.85**
Discrimination as % total (D/R)	**56.31**	**56.64**
Total	100	100

Table 8. Full Decomposition of Gender Earnings Gap Accounting for Selectivity

Earnings differences due to within-sector differences attributable to		
Characteristics $\sum_{j=1}^{3} \overline{p}_j^* (\overline{x}_{mj} - \overline{x}_{fj}) \beta_j$	0.071	29.4%
Deviation in male returns $\sum_{j=1}^{3} \overline{p}_j^* \overline{x}_{mj} (\beta_{mj} - \beta_j)$	-0.552	-229.9%
Deviation in female returns $\sum_{j=1}^{3} \overline{p}_j^* \overline{x}_{fj} (\beta_j - \beta_{fj})$	-0.087	-36.4%
Sub-total	-0.569	
Earnings differences due to between-sectoral location attributable to		
Characteristics $\sum_{j=1}^{3} \overline{W}_{mj} (\overline{p}_{mj}^* - \overline{p}_j^*) + \sum_{j=1}^{3} \overline{W}_{fj} (\overline{p}_j^* - \overline{p}_{fj}^*)$	0.057	23.7%
Deviation in effect of characteristics on male sectoral location $\sum_{j=1}^{3} \overline{W}_{mj} (\overline{p}_{mj} - \overline{p}_{mj}^*)$	0.729	303.4%
Deviation in effect of characteristics on female sectoral location $\sum_{j=1}^{3} \overline{W}_{fj} (\overline{p}_{fj}^* - \overline{p}_{fj})$	0.023	9.8%
Sub-total	0.809	
Total	**0.240**	100%

These decompositions stem from earnings regressions that include actual experience variables (tenure and previous experience) and limited LFAVs.

References

Adair L., Guilkey D., Bisgrove E. and Gultiano S. (2002), "Effect of Childbearing on Filipino Women's Work Hours", *Journal of Population Economics*, 15(4), pp. 625-645.

Agesa R.U. (1999), "The Urban Gender Wage Gap in an African Country: Findings from Kenya", *Canadian Journal of Development Studies*, 20(1), pp. 59-76.

Altonji J.G. and Blank R.M. (1999), "Race and Gender in the Labor Market", in Ashenfelter O.C. and Card D. (eds), *Handbook of Labor Economics*, vol. 3C, Elsevier, chapter 48, pp. 3143-3258.

Antecol H. and Bedard K. (2004), "The Racial Wage Gap: The Importance of Labor Force Attachment Differences across Black, Mexican, and White Men", *The Journal of Human Resources*, 34(2), pp. 564-583.

Antoine Ph., Bocquier Ph., Maminirina T. and Razafindratsima N. (2004), "Collection of Biographical Data in Antananarivo: the *Biomad98* Survey", *Interstat* 28, April, pp. 83-99.

Antoine Ph., Bocquier Ph., Razafindratsima N. and Roubaud F. (2000), *Biographies de trois générations dans l'agglomération d'Antananarivo*, Ceped, collection Documents et manuels n°11, Paris.

Appleton S., Hoddinott J. and Krishnan P. (1999), "The Gender Wage Gap in Three African Countries", *Economic Development and Cultural Change*, 47(2), pp. 289-312.

Armitage J. and Sabot R. (1991), "Discrimination in East African's Urban Labor Market", in Birdsall and Sabot (eds), *Unfair Advantage: Labor Market Discrimination in Developing Countries*, Washington D.C., The World Bank.

Barron J.M., Berger M.C. and Black D.A. (1999), "Do Workers Pay for On-the-Job Training", *The Journal of Human Resources*, 34(2), pp. 235-252.

Behrman J.R. (1999), "Labor Markets in Developing Countries", Chapter 43 in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, Elsevier Science Publishers BV, Netherlands.

Bennell P. (1996), "Rates of Return to Education: Does the Conventional Pattern Prevail in Sub-Saharan Africa?", *World Development*, 24(1), pp. 183-199.

Blau F. and Kahn L., (2000), "Gender Differences in Pay", *Journal of Economic Perspectives*, 14, pp. 75-99.

Blinder A. (1973), "Wage Discrimination: Reduced Form and Structural Estimates", *The Journal of Human Resources*, 8(4), pp. 436-455.

Booth A., M. Francesconi, and J. Frank (2003), "A Sticky Floors Model of Promotion, Pay and Gender", *European Economic Review*, 47, pp. 295-322.

Bourdieu P. (1998), *La domination masculine*, Paris, Seuil, coll. Liber, 134 p.

Bourguignon F., Fournier M. and Gurgand M. (2004), "Selection Bias Corrections Based on the Multinomial Logit Model: Monte-Carlo Comparisons", DELTA Working Paper, September.

Brown R.S., Moon M. and Zoloth B.S. (1980), "Incorporating Occupational Attainment in Studies of Male-Female Earnings Differentials", *The Journal of Human Resources*, 15(1), pp. 3-28.

Card D. (1999), "The Causal Effect of Education on Earnings", in Ashenfelter O. and Card D. (eds), *Handbook of Labor Economics*, vol. 3A, Elsevier Science Publishers BV, Netherlands, pp. 1801-1863.

Cling J.-P., Razafindrakoto M., Roubaud F. (2003), *New International Poverty Reduction Strategies*, London/New York, Routledge.

Cling J.-P., Razafindrakoto M., Roubaud F. (2005), "Export Processing Zones in Madagascar: a Success Story under Threat?", *World Development*, 33(5), pp.785-803.

Cogneau D. (1999), "La formation du revenu des ménages à Antananarivo : une microsimulation en équilibre général pour la fin du siècle", *Economie de Madagascar*, 4, pp. 131-155.

Cohen B. and House W.J. (1993), "Women's Urban Labour Market Status in Developing Countries: How Well Do They Fare in Khartoum, Sudan?", *Journal of Development Studies*, 29(3), pp. 461-83.

Courgeau D. and Lelièvre E. (1992), *Event History Analysis in Demography,* Clarendon Press, Oxford University Press.

Destré G. and Nordman C. (2002), "The Impacts of Informal Training on Earnings: Evidence from French, Moroccan and Tunisian Employer-Employee Matched Data", *SSRN Electronic Library, Human Capital Journal*. Published in French in *L'Actualité Économique*, 78(2), pp. 179-205.

Dougherty C. (2003), "Why is the Rate of Return to Schooling Higher for Women Than for Men?", Centre for Economic Performance Discussion Paper, #581, August.

Glewwe P. (1990), "Schooling, Skills and the Return to Education: an Econometric Exploration Using Data from Ghana", *Living Standards Measurement Working Paper*, n°76, Washington D.C., The World Bank.

Glick P. and Sahn D. E. (1997), "Gender and Education Impacts on Employment and Earnings in West Africa: Evidence from Guinea", *Economic Development and Cultural Change*, 45(4), pp. 793-823.

Goux D. and Maurin É. (1999), "Persistence of Interindustry Wage Differentials: A Reexamination Using Matched Worker-Firm Panel Data", *Journal of Labor Economics*, 17(3), pp. 492-533.

GRAB (Groupe de Réflexion sur l'Approche Biographique) (1999), *Biographie d'enquêtes: bilan de 14 collectes biographiques*, Collection « Méthodes et savoirs », PUF diffusion, Paris.

Hoddinott J. (1996), "Wages and Unemployment in an Urban African Labour Market", *The Economic Journal*, 106, pp. 1610-1626.

Hussmann R., Mehran F. and Verma V. (1990), *Surveys of Economically Active Population, Employment, Unemployment and Underemployment. An ILO Manual on Concepts and Methods.* Geneva.

Isemonger A. G. and Roberts N. J. (1999), "Post-entry Gender Discrimination in the South African Labour Market", *Journal for Studies in Economics and Econometrics*, 23 (2), pp. 1-25.

Kim M.-K. and Polachek S.W. (1994), "Panel Estimates of Male-Female Earnings Functions", *The Journal of Human Resources*, 29(2), pp. 406-428.

Lee L.-F. (1983), "Generalized Econometric Models with Selectivity", *Econometrica*, 51(2), pp. 507-512.

Light A. and Ureta M. (1995), "Early Career Work Experience and Gender Wage Differentials", *Journal of Labor Economics*, 13(1), pp. 121-154.

MADIO (1998), *L'emploi, le chômage et les conditions d'activité dans l'agglomération d'Antananarivo*, Enquête 1-2-3, Premiers résultats de la phase 1 sur l'emploi, Instat, Madagascar.

MADIO (1999), *Trajectoires biographiques dans l'agglomération d'Antananarivo : Emploi, Famille*, Logement, premiers résultats de l'enquête Biomad 98.

Meurs D. and Ponthieux S. (2000), "Une mesure de la discrimination dans l'écart de salaires en hommes et femmes", *Économie et Statistiques*, 337-338, pp. 135-158.

Milne W. and Neitzert M. (1994), "Kenya", in Horton S., Kanbur R. and Mazundar D. (eds), *Labor Markets in an Era of Adjustment: Issues Papers*, 1, Washington D.C, EDI Development Studies, World Bank, pp. 405-457.

Mincer J. (1993), *Schooling, Experience, and Earnings*, New York: Gregg Revivals, 1993 (first edition: 1974).

Mincer J. and Jovanovic B. (1981), "Labor Mobility and Wages", in S. Rosen (ed), *Studies in Labor Markets*, pp. 21-64, Chicago, University of Chicago Press.

Mincer J. and Ofek H. (1982), "Interrupted Work Careers: Depreciation and Restoration of Human Capital", *The Journal of Human Resources*, 17, pp.1-23.

Myck M. and Paull G. (2004), "The Role of Employment Experience in Explaining the Gender Wage Gap", Institute for Fiscal Studies Working Papers, W04/16, July.

Neuman S. and Oaxaca R.L. (2004), "Wage Decompositions with Selectivity-Corrected Wage Equations: A Methodological Note", *The Journal of Economic Inequality*, 2(1), April, pp. 3-10.

Neumark D. (1988), "Employers' Discriminatory Behavior and the Estimation of Wage Discrimination", *The Journal of Human Resources*, 23, pp. 279-295.

Nicita A. and Razzaz S. (2003), "Who Benefits and How Much? How Gender Affects Welfare Impacts of a Booming Textile Industry", World Bank Policy Research Working Paper, #3029, April.

Nordman (2003), “Formal and Informal Training: a Critical Review of its Economic Features and Measures”, SKOPE Research Paper, Department of Educational Studies, University of Oxford.

Nordman C. (2004), “Discrimination salariale, capital humain et structure des tâches selon le genre : l'apport de données liées employeurs-employés au Maroc et en Tunisie”, in Maruani M., Meulders D., Sofer C. (eds), *Marché du travail et genre : Maghreb-Europe,* Chapter 2, Brussels Economic Series, Editions du Dulbéa, Brussels.

Nordman C. and Hayward G. (2004), “The Role of Skill Usage, Tasks and Workstation in Mediating the Benefits of Formal and Informal Training: Evidence from British Workers”, SKOPE Research Paper, Department of Educational Studies, University of Oxford.

Oaxaca R. (1973), “Male-Female Wage Differentials in Urban Labor Markets”, *International Economic Review*, 14(3), pp. 693-709.

Oaxaca R. and Ransom M. R. (1994), “On Discrimination and the Decomposition of Wage Differentials”, *Journal of Econometrics*, 61(1), pp. 5-22.

O'Neill J. and S. Polachek (1993), “Why the Gender Gap in Wages Narrowed in the 1980s”, *Journal of Labor Economics*, 11(1), pp. 205-228.

Rakotomanana F., Ramilison E. and Roubaud F. (2003), “The Creation of an Annual Employment Survey in Madagascar. An Example for Sub-saharan Africa”, *InterStat* 27, September, pp. 35-57.

Razafindrakoto M. and Roubaud F. (1999), “La dynamique du marché du travail dans l’agglomération d’Antananarivo entre 1995 et 1999 : la croissance économique profite-t-elle aux ménages ?”, *Economie de Madagascar*, 4, pp.103-137.

Razafindrakoto M. and Roubaud F. (2003), “Two Original Poverty Monitoring Tools: the *1-2-3 Surveys* and the Rural Observatories”, in Cling J.-P., Razafindrakoto M., Roubaud F. (eds), *New International Poverty Reduction Strategies*, London/New York, Routledge, pp. 313-339.

Reimers C.W. (1983), “Labour Market Discrimination Against Hispanic and Black Men”, *The Review of Economics and Statistics*, 65(4), pp. 570-579.

Roubaud F., ed. (2000), *Le projet MADIO à Madagascar : l’information statistique au service du débat démocratique sur la politique économique*, Stateco, INSEE, n°95/96/97, France.

Roubaud F., ed. (2002), *Madagascar après la tourmente : regards sur dix ans de transitions politique et économique*, dossier spécial, *Afrique Contemporaine*, n° 202/203, April-September, pp. 3-163.

Sahn D.E. and Alderman H. (1988), “The Effects of Human Capital on Wages and the Determinants of Labor Supply in a Developing Country”, *Journal of Development Economics*, 29, pp. 157-183.

Sandell S.H. and Shapiro D. (1980), “Work Expectations, Human Capital Accumulation, and the Wages of Young Women”, *The Journal of Human Resources*, 15(3) pp. 335-353.

Schultz T. P. (2004), “Evidence of Returns to Schooling in Africa from Household Surveys: Monitoring and Restructuring the Market for Education”, *Journal of African Economies*, 13, AERC Supplement, pp. ii95-ii148.

Siphambe H.K. and Thokweng-Bakwena M. (2001), “The Wage Gap Between Men and Women in Botswana's Formal Labour Market”, *Journal of African Economies*, 10(2), pp. 127-142.

Stanley T.D. and Jarrell S.B. (1998), “Gender Wage Discrimination Bias? A Meta-Regression Analysis”, *The Journal of Human Resources*, 33(4), pp. 947-973.

Weichselbaumer D. and Winter-Ebmer R. (2005), “A Meta-Analysis of the International Gender Wage Gap”, *Journal of Economic Surveys*, 19(3), pp. 479-511.

Wright R. and J. Ermisch (1991), “Gender Discrimination in the British Labour Market: a Reassessment”, *Economic Journal*, 406(101), pp. 508-522.

Appendix
Not for publication

Probit Estimates of Males and Females' Employment Participation

	Overall sample	Males	Females
	(1)	(2)	(3)
Individual characteristics			
Sex	-0.1468	_	_
	(1.36)		
Age	0.1174***	0.2605***	0.0735
	(2.73)	(3.58)	(1.36)
$(\text{Age})^2$	-0.0019***	-0.0041***	-0.0012*
	(3.38)	(4.36)	(1.80)
Catholic	0.1130	0.3442**	0.0374
	(1.37)	(2.18)	(0.36)
Merina	0.2672**	0.3793**	0.1539
	(2.41)	(2.22)	(1.06)
Household head	0.7576***	0.7483***	0.6433***
	(5.24)	(3.27)	(3.11)
Head’s spouse	-0.2467	_	0.0748
	(1.60)		(0.41)
Head’s children	0.1423	-0.1353	0.3708*
	(1.00)	(0.61)	(1.88)
Head’s parent	-0.6594	_	-0.8697
	(0.84)		(0.97)
Married	0.4470***	0.9837***	-0.0365
	(3.29)	(4.04)	(0.21)
Married before first employment	0.4351***	0.2167	0.4455***
	(3.86)	(0.73)	(3.84)
Years of completed schooling	0.0337***	0.0294	0.0436***
	(2.88)	(1.13)	(3.02)
Total actual labour market experience	0.0490***	0.0702***	0.0465***
	(10.29)	(4.13)	(9.24)
Household’s characteristics			
Household’s income per capita	-0.0006**	-0.0010**	-0.0005*
	(2.22)	(2.34)	(1.66)
Proxy of material wealth	0.0415**	0.1383***	0.0122
	(2.12)	(3.81)	(0.61)
Inverse dependency ratio	0.2176	0.5852	-0.0273
	(0.99)	(1.43)	(0.10)
Number of children between 0 and 4 years old	0.0662	0.3429***	-0.1002
	(1.07)	(3.01)	(1.27)
Number of children between 5 and 9 years old	0.1231**	0.1043	0.0676
	(2.07)	(0.98)	(0.94)
Number of children between 10 and 14 years old	0.1672***	0.2172**	0.0967
	(2.74)	(2.14)	(1.29)
Number of children below 15 for married individuals	-0.1698***	-0.3096***	-0.0616
	(3.16)	(3.06)	(0.95)
Father’s education	-0.0200*	-0.0334	-0.0174
	(1.75)	(1.63)	(1.23)

Spouse never went to school	-0.1200	-0.1079	-0.1101
	(1.20)	(0.50)	(0.95)
Spouse has higher education level	0.1484	0.4882*	0.0782
	(1.41)	(1.92)	(0.64)
Spouse is Catholic	-0.0181	0.0973	-0.0403
	(0.20)	(0.51)	(0.37)
Spouse is another religion	0.2074	0.8167*	0.0819
	(1.06)	(1.95)	(0.33)
Spouse is Merina	0.0885	-0.3043	0.1352
	(0.76)	(1.42)	(0.95)
Spouse is other ethnic group	-0.1319	-0.6198**	-0.0013
	(0.81)	(2.19)	(0.01)
Housing characteristics			
Tenancy	0.0795	-0.1551	0.1837
	(0.69)	(0.69)	(1.21)
Individual is harboured	-0.1241	-0.3761*	-0.0736
	(1.28)	(1.69)	(0.66)
Receives electricity	0.4725***	0.7843***	0.3906***
	(4.70)	(3.19)	(3.38)
Constant	-2.5797***	-5.1057***	-1.7665*
	(3.14)	(3.61)	(1.70)
Pseudo R-squared	0.26	0.38	0.15
Log pseudo-likelihood	-858.30	-214.38	-613.57
Observations	2334	1149	1181

Robust z statistics in brackets. ***, ** and * mean respectively significant at the 1%, 5% and 10% levels.

Maximum Likelihood Estimates of Multinomial Logit Sectoral Choice Models

	Males			Females		
	Public wage employment	**Formal private wage employment**	**Self-employed or informal sector**	**Public wage employment**	**Formal private wage employment**	**Self-employed or informal sector**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Individual characteristics						
Age	0.9744***	0.5956***	0.6108***	0.7464***	0.2095*	0.1239
	(4.90)	(3.49)	(3.45)	(4.01)	(1.71)	(1.15)
$(Age)^2$	-0.0135***	-0.0100***	-0.0102***	-0.0095***	-0.0041**	-0.0023*
	(5.36)	(4.41)	(4.37)	(4.16)	(2.57)	(1.67)
Catholic	0.4527	0.6824*	0.4044	0.1427	0.1059	0.2652
	(1.21)	(1.94)	(1.11)	(0.42)	(0.44)	(1.27)
Merina	0.7332*	0.8965**	1.0522**	0.8451**	0.6726**	0.4227
	(1.66)	(2.14)	(2.42)	(2.02)	(2.01)	(1.36)
Household head	2.4304***	1.4319***	1.6952***	1.9504***	1.4522***	0.0762
	(3.81)	(2.60)	(2.88)	(2.97)	(2.89)	(0.17)
Head’s spouse	-7.1873***	19.0399	19.6247***	0.8827	0.6051	-0.4705
	(6.69)	(.)	(18.01)	(1.39)	(1.37)	(1.21)
Head’s children	-0.6695	-0.6006	-0.0554	1.4667**	0.9579**	-0.0863
	(0.97)	(1.16)	(0.10)	(2.28)	(2.11)	(0.22)
Head’s parent	0.1268	28.6848***	-1.3957	-33.8640***	0.8353	-38.3158***
	(0.12)	(20.86)	(1.44)	(22.33)	(0.47)	(27.93)
Married	2.6475***	2.1795***	1.9895***	1.1056**	0.1133	-0.2880
	(4.00)	(3.44)	(3.08)	(2.09)	(0.30)	(0.79)
Married before first employment	0.7460	0.4140	1.0479	1.1285***	1.1959***	0.9906***
	(0.98)	(0.55)	(1.36)	(3.52)	(4.32)	(3.56)
Years of completed schooling	0.1619**	0.0621	-0.0198	0.4023***	0.0967***	-0.0554*
	(2.52)	(1.04)	(0.31)	(8.11)	(2.79)	(1.70)
Total actual labour market experience	0.1885***	0.1839***	0.1993***	0.1687***	0.1155***	0.0888***
	(4.49)	(4.51)	(4.74)	(9.19)	(8.98)	(7.87)

Household characteristics						
Household's income per capita	-0.0015	-0.0028***	-0.0051*	-0.0012*	-0.0010	-0.0010
	(1.47)	(2.72)	(1.83)	(1.77)	(1.36)	(1.00)
Proxy of material wealth	0.2941***	0.3792***	0.1362	-0.0541	0.0063	-0.0611
	(3.26)	(4.32)	(1.42)	(0.92)	(0.17)	(1.29)
Inverse dependency ratio	2.9853***	3.5221***	2.5231**	-0.7354	0.4013	0.3411
	(3.04)	(3.74)	(2.37)	(0.94)	(0.65)	(0.59)
Number of children between 0 and 4 years old	0.8604***	0.6164**	0.6053**	0.0829	-0.4258**	-0.1440
	(2.88)	(2.21)	(2.21)	(0.32)	(2.16)	(0.92)
Number of children between 5 and 9 years old	0.2478	-0.1048	-0.0438	0.2603	-0.0475	0.0988
	(0.98)	(0.40)	(0.17)	(1.20)	(0.27)	(0.67)
Number of children between 10 and 14 years old	0.5405**	0.3116	0.2631	0.6129***	-0.0021	0.1218
	(2.11)	(1.19)	(1.08)	(2.89)	(0.01)	(0.79)
Number of children below 15 for married individuals	-0.7852***	-0.5043**	-0.6015***	-0.5088**	-0.0420	0.0446
	(3.28)	(2.08)	(2.60)	(2.46)	(0.26)	(0.35)
Father's education	-0.1063**	-0.0890*	-0.0774	-0.0370	-0.0350	-0.0661**
	(2.03)	(1.82)	(1.43)	(0.94)	(1.17)	(2.06)
Spouse never went to school	-0.4629	-0.4056	-0.3273	-1.1843**	-0.5290*	-0.1039
	(0.82)	(0.74)	(0.59)	(2.45)	(1.93)	(0.45)
Spouse has higher education level	0.8321	0.7203	1.0422*	-0.1140	-0.1692	-0.4559
	(1.34)	(1.18)	(1.66)	(0.37)	(0.66)	(1.59)
Spouse is Catholic	0.8011	0.2502	0.5214	0.0003	-0.2309	-0.3415
	(1.53)	(0.49)	(1.01)	(0.00)	(0.90)	(1.54)
Spouse is another religion	0.9757	1.3169*	1.2591	1.1702*	0.2425	0.2388
	(1.13)	(1.66)	(1.53)	(1.86)	(0.40)	(0.47)
Spouse is Merina	-1.9433***	-0.6928	-1.4324***	-0.4684	0.1051	0.1118
	(3.49)	(1.30)	(2.69)	(1.15)	(0.31)	(0.36)
Spouse is other ethnic group	-1.8681***	-0.8241	-1.7620**	-0.4795	-0.1446	-0.0202
	(2.68)	(1.23)	(2.30)	(0.89)	(0.33)	(0.05)
Housing characteristics						
Tenancy	-0.1028	0.0233	-0.1223	0.8565**	0.6013*	0.1984

	(0.17)	(0.04)	(0.20)	(1.97)	(1.77)	(0.59)
Individual is harboured	-0.1049	-0.2979	-0.2081	0.0467	-0.5235**	-0.3283
	(0.18)	(0.52)	(0.35)	(0.15)	(2.06)	(1.42)
Receives electricity	1.5921***	1.7305***	2.2162***	0.0865	-0.2209	0.6270***
	(2.69)	(3.07)	(4.02)	(0.21)	(0.81)	(2.72)
Constant	-23.4134***	-12.5493***	-11.5457***	-22.9105***	-4.7185**	-1.1551
	(5.98)	(3.87)	(3.42)	(6.00)	(2.01)	(0.54)
Pseudo R-squared		0.21			0.25	
Log pseudo-likelihood		-1148.24			-1169.71	
Observations		1153			1181	

The reference group is non-participation in paid employment. Robust z statistics are in brackets. ***, ** and * mean respectively significant at the 1%, 5% and 10% level.